

CHAPTER I

INTRODUCTION

The first chapter deals with Background of Study, Reason for Choosing Topic, Statement of the Problems, Objective of the Study, Hypotheses of the Study, Limitation of the Study, Significance of the study, Definition of Key Terms, and Organization of the study.

1.1. Background of the Study

The test is kind of a method that is used to evaluate or to measure the learning and teaching process. There are two kinds of test, it can be in the form of test and non test. Most teachers use a written test to make an assessment. The purposes of the test are to measure the students' ability when they did the test items, to know how far the students mastering the material that was delivered by teachers, and also to help the teacher knowing how many students can do the test items correctly. The test becomes an important part in learning and teaching process for accumulating event. Hughes (2003:13) cited in Lilia Indriani (2013) said "Test is expected to measure the students' accomplishment and level of successful learning and teaching program". By testing, we can measure the students' ability and their knowledge. Not only that, through testing their progress can be known by themselves, and teachers can adjust the way they teach everyday. To sum up, to measure the extent to which students' ability to master the material, tests are the right tool to help teachers find out. In addition, By using tests, the strengths and weaknesses of students can also be identified, this is to

ensure what lessons are still needed. According to Wilmar Tinambunan in his book entitled *Evaluation of Students Achievement* (1998) cited on Noviyanti (2011) four kinds of the test like placement test, formative test, diagnostic test, and summative. Placement test means a test that is given by the school to the students when they entering an educational institution, to decide and to place, whether a student is ready in a specific subject or not. Formative test is used to measure students' ability during the learning process, and is designed to give students the opportunity to show that they have understood the material. Diagnostic test called as a test conducted at the beginning of students entering a new phase of education, this test is used to diagnose what they have not and already know about the topics that students will be taught in the upcoming lessons. Summative test is usually done at the end of the academic year to assess how much the material that students have learned overall.

A good test has five criteria, such as practicality, reliability, validity, authenticity, and washback. Practicality means cheap, right time, simple to do, specific in the scoring process, and time efficient. Reliability means that the test has to consistence and dependable, when two tests are given by the teacher at different times the results will be the same or similar. Validity means the test is more complex of effective test criterion. Gronlund (1998) stated to find out whether the result made from the conclusion of assessment are already appropriate, have meaningful and useful for research. Authenticity is a little difficult concept to define, especially within the art and science of evaluating and designing tests. Bachman and Palmer, (1996) stated that authenticity as "the correspondence's

level of the characteristics of a selected language test task with the features's type of a target language". And the last criteria of a good test is wash back, wash back is the effects the tests have on instruction in terms of how students prepare for the test.

In addition, good quality of a test is also depends on the quality of test items. That is why the analysis of test items is an important thing to do. To analyze the test items, such as validity, reliability, level difficulty, discriminating power and distractor of test items. Thus, to identify or solve the most appropriate test questions and ensure that new tests are designed to meet student needs and truly reflect learning and meet good test requirements, test analysis is expected to help teachers.

From the explanation above, the writer wants to do the analysis of English summative test item quality for senior high school students' academic year 2018/2019, which has been made by teacher or the government of Semarang City to find out whether or not the test items are functioning properly. In this study, the writer focuses on multiple choice tests.

1.2. Reasons for Choosing the Topic

There are some reasons why the writer chooses the topic:

1. The quality of English summative test items in academic year 2018/2019 is unknown.
2. Research about quality of test items is still rarely done, especially of the English summative test for 11th grade in academic year 2018/2019.

3. In common problem is when a quiz or test is held, there are some questions that only a few students can answer correctly. Therefore, an analysis of the test item needs to be done, to find out the reason why the question is difficult to answer.
4. As we know, lattices and questions are an inseparable package. If there is an imbalance, it needs to be taken seriously. Why did it happen like that? Therefore, item analysis is needed.
5. Sometimes, there are some test items that are not in accordance with the syllabus or material that has been taught. It makes students unable to answer the questions. Therefore, an item analysis is needed to make the quality of the questions made by the teacher better.

1.3. Question of the Study

Related to the background of the study, the question of the study: "How is the quality of English summative test Item for senior high school students in Semarang academic year 2018/2019 according to the aspect of validity, reliability, level of difficulty, discriminating power, and aspect of distractor?"

1.4. Objective of the Study

The objective of this study is to describe the quality of English summative test items for senior high school students in Semarang academic year 2018/2019 according to the aspect of validity, reliability, level of difficulty, discriminating power, and aspect of distractor.

1.5. Hypothesis of the Study

Sarantakos and Shalini Prasadstated (2011) that a hypothesis is a tentative explanation of the research problem, a possible outcome of the research, or an educated guess about the research outcome.

Based of the statement of the problem, the hypotheses of this study are as follows:

H1 : The quality of English summative test items for senior high school students in Semarang, in the academic year 2018/2019 is good according to the aspect of validity, reliability, level of difficulty, discriminating power, and aspect of distractor.

H0 : The quality of English summative test items for senior high school students in Semarang, in the academic year 2018/2019 is not good according to the aspect of validity, reliability, level of difficulty, discriminating power, and aspect of distractor.

1.6. Limitation of the Study

This analysis only covers the English summative test for senior high school students in 11th grade, carried out in academic year 2018/2019 administered by the government of Semarang City, and focusing only on the multiple choice items.

1.7. Significance of the Study

The results of this study are expected can give the following benefits :

1. Pedagogical significance

One of competencies that must be mastered is pedagogical competence. Where the teacher is not only be able in making or compile evaluation tools to find out achievement of student learning outcomes but can also evaluate the result of it. Therefore, the results of this study are expected to describe whether the evaluation that has been prepared can function as a measure of learning outcomes that has good quality.

2. Theoretical significance

- a. The results of this study are expected to contribute to the development of education, especially in the field of learning outcome assessment.
- b. As a reference for the next research.

3. Practical significances

a. For the Teacher or Test Maker

This research is expected to give the benefits for input and as a suggestion that can be used as a reference in making English summative test items in the next period.

b. For School

This research is expected to be useful as a reference for improving the quality of the next exam test items. Therefore, they can measure student's learning outcomes well and can assess teacher success in learning activities.

c. For Researchers

This research is expected to open researchers' discourse about new things and can be useful to increase knowledge and skills in applying test item analysis of English subjects.

1.8. Definition of the Key Term

Summative test

According to Jack and R. Keyworth (2018) summative test is the test that happened at the last of an academic year of school, usually it is used to measure the students' ability annually. In line, that summative test is a result of learning test carried out after a set of teaching program given completely.

1.9. Organization of the Study

This organization of this proposal is arranged into three chapters.

Chapter I consists of the Background of Study, Reason for Choosing Topic, Statement of the Problems, Objective of the Study, Hypotheses of the Study, Limitation of the Study, Significance of the study, Definition of Key Terms, and Organization of the study.

Chapter II deals with Testing, Multiple Choice Question, Validity, Reliability, Item Analysis, ITEMAN and Previous Studies.

Chapter III presents methodology of Research Design, Subject of the Study, Data for the Study (Type of Data), Instrument, Technique for Collecting Data, Data Collecting Procedure, Data Analysis, and Time Schedule.

Chapter IV describes Finding and Discussion which consist of Description of School's Profile, Description of the Participants, Validity and Reliability of the

Instrument, Analysis of aspect validity, reliability, level of difficulty, discriminating power, distractor, and Discussion of the Research Findings.

Chapter V is about the Conclusion of this Study and some suggestions based on the result of this study.