CHAPTER I

INTRODUCTION

This chapter will discuss the background of the research, research questions, research purposes, significant of research, research framework and previous research.

A. Background

Why do some students assume that English test is difficult? Is it because of their less knowledge? Is it because of the teacher's complicated explanation while teaching? Or is it because of the quality of the test? According to Brown (2006), the content in multiple tests must be tested with the two language assessment principle: practicality and reliability. Hughes (2003), stated that the test content and testing techniques must have the same objective as the course; if it is not balance, it will be harmful in backwash. So, the quality of the test must be reliable and practical because it will affect students' scores and it also will measure students' progress in the class and teacher's ability in making a test before conducting an examination. So, how can the teacher make a good test for their students?

A good test is equipment that gives an accurate measurement of students' capability in a particular domain. Brown (2006) defines a test as a measurement of someone's ability, knowledge, or performance in certain things. Hughes (2003) interprets a test as a method in which the information about someone's skills can be assembled and it is just one form of assessment and another way will often be more appropriate. Furthermore, Harmer (2010) said that a good test is those who can be done as they are designed do and convinces people to take and mark them that they work.

In making a test, teacher should know about the kinds of assessment that would be designed because every type of assessment has a different way to prepare. Most of our classroom assessments are formative and summative assessment. Formative assessment is intended to improve learning by providing

feedback to students before instructors issue performance evaluations. Formative assessments identify learners' strengths and weaknesses throughout the learning cycle and, as a result, aim to improve future performance. Meanwhile, summative assessment (evaluation/grading) is used to determine readiness for progression by providing performance evaluations. Summative assessment, as the name implies, occurs at the end of an educational activity or learning cycle (such as mid-term test, final test, etc.) and is intended to evaluate the learner's overall performance (knowledge, skills sets, etc.). Summative evaluations are serve as the foundation for grading (Wright, 2012).

Besides knowing the type of test, in creating a test, we are as teachers have to recognize the principle of language assessment. Brown (2006) stated in his book that a test is said to be productive when it can be answered by some questions. Those questions will help in identifying five important measurements for testing a test. The five principals are practicality, reliability, validity, authenticity, and washback. Besides the principals, the quality of test also has to be considered such as whether the teacher can analyze the difficulty of their trial by item facility, item discrimination, and distractor efficiency (Brown, 2006).

English teachers should be able to assess and evaluate their students' learning achievement. A test is a subset of assessment. Most English teachers conduct a test or examination to measure their students' achievement in the learning process. A test might measure linguistic competence or particular skills. A test can be inform of written or oral test (Gultom, 2016). The overall goal of assessment is to improve students learning. Assessment should always be viewed as information to improve students' achievement (Meidasari, 2017). The significant impact of the assessment, such as conducting a test for English language learners, is to know how far they have learned, whether it is good or not, and whether they improve their knowledge in mastering the English language.

Based on the researcher's observation, when she carried out pre-teacher activities in some schools, she knows that one of the teachers there had held a review exam on English subjects. When the researcher asked to help him check the students' answers, the researcher paid attention to each item of the questions,

and it looked comfortable for the students. Apparently, the problem of the test is that it does not have variation in its item. It is realized that in measuring students' ability, the teacher should pay attention to each item on the test given to students. Therefore, the teacher must be careful in making a test because it is not easy and some principles must be considered. Hughes (2003) stated that a test is a tool for teachers to know how students' ability in mastering some lesson. So, teachers should know the best way to make a test while conducting an examination in for their students. As Brown (2006) stated, when we want to start designing a test, we should know the basis of our approach to design a test. It can measure students' abilities and it can be a great skill that teachers possess to be a professional teacher.

There are several research regarding the Language Assessment. Many research identified the importance of language assessment principal in English tests. One, a research by Indriani (2004) investigates a language assessment principal in a review test. The result shows that the review test is still not suitable with the language assessment principle because there are some mistakes invalidity aspects. Meanwhile, this research centers on analyzing items in the English multiple-choice test. Due to the reasons, this study focuses on investigating the quality of the test through item analysis that is usually made by teachers before conducting an examination (summative test).

B. Research Questions

1. How is the item's quality of the English summative multiple-choice test question?
2. How are the teachers' steps to design the English summative multiple-choice test item?

C. Research Purposes

1. To identify the item's quality of the multiple-choice test questions.
2. To describe the teacher's steps in designing a multiple-choice test item.

D. The significances of research

This study is considered essential due to information about how vital the quality in multiple-choice tests except for the language assessment principle and expected to give the teachers information to design a multiple-choice test well. Practically, this research provides significance to teachers. This research is expected to show how important the quality of a test, especially in multiple-choice tests. Furthermore, this study is expected to help the teachers to design a multiple-choice test well. The result of this research is expected to make the teachers aware when they want to give their students multiple tests because it will measure the students' progress in the teaching and learning process.

Theoretically, this study is expected to be beneficial for test designers as an evaluation to design the best multiple-choice test.

E. Research Framework

One thing to support the teaching and learning process is an assessment. Teachers will know about students' abilities by testing students' knowledge and assessing their performance. Assessment serves multiple purposes such as providing information about students' learning and progress, teaching quality and program, and institutional accountability (Opre, 2015). Assessment is an integral part of the instructional process. It determines where learners are at present and what level they have achieved, gives learners feedback on their learning, diagnoses learners' needs for further development, enabling the planning of curricula, materials, and activities (Alderson, 2005 in Sahinkarakas, 2012). According to Hughes (2003), a test is the only way to gather information about people's language ability. Information means that teachers can know everything new from someone's test, such as knowledge, capability, and others.

In general, assessment and evaluation should be carried out by every EFL teacher. As stated before, most teachers conduct a test or examination to measure a student's achievement in EFL teaching and learning. English teaching should be able to assess and evaluate their students' learning achievement (Gultom, 2016).

One of the assessment devices is a test. Tests are a subset of assessment. Although they are not the only form of assessment, tests can be useful devices to assess students (Brown, 2004 in Hudaya, 2017). While the teaching process, teachers positively assess students' performance when they are learning something. Brown (2006) stated, "During practice activities, teachers are indeed observing students' performance and making a various evaluation of each learner".

Multiple-choice tests have been used extensively for many years for assessment purposes (Roberts, 2006). Many teachers use multiple choice tests to be an assessment in grading students' capability. It becomes an alternative test because of the natural way in design and it can shorten time while assessing. Multiple-choice items, which may appear to be the most straightforward kind of items to construct, are complicated to design correctly (Brown, 2006).

Multiple-choice test has great advantage of being easy to mark. One problem with multiple-choice tests lies in the choice of distractors that are the three incorrect answers. Sometimes, it may not be difficult to make one obvious the distractors. However, it becomes less easy to come up with three items that will all sort out those students who know how this piece of language works from the ones who don't. In other words, there is a danger that will either distract too many students (even those who should get the question right) or too few (in which case the question has not done its job of differentiating students) (Harmer, 2010).

Brown (2006) stated in his book that a test is called useful when some questions can be answered. Those questions will help in identifying five necessary measurements for testing a test. The five principals are practicality, reliability, validity, authenticity, and washback. Besides the principals, the quality of test also has to be considered such as whether the teacher can analyze the difficulty of their trial by item facility, item discrimination, and distractor efficiency.

F.   Previous Research

There are several research regarding Language Assessment. First, a research by Paramatha (2017) with the title "The analysis of the multiple-choice test

quality for reading III class in English Education Department, Universitas Pendidikan Ganesha Bali, Indonesia". It analyzes the quality of the 40-item multiple-choice test. The method of the research is the quantitative method because the data were analyzed by conducting Classical Test Theory analysis with the assistance of jMetrik software. The result shows that the test was consistent, but it was easy, and only 40% of a total item are eligible to be used.

Second, a research by Grumilah & Aji (2006) with the title "Informal assessment for language skills: the learners' perspective". It is about students' perspective toward the use of informal assessment. The research shows that students' achievement in the teaching and learning process could not be measured accurately through a formal assessment only. An informal assessment is needed too in measuring student's achievement. This research uses quantitative method. The technique of collecting the data is through questionnaire. They choose 211 students randomly selected from 600 students of 20 English language skills classes in PBI Universitas Ahmad Dahlan during the second semester of the academic year 2013-2014. The finding shows that students' perspective is proper in the informal assessment.

Third, a research by Sabrina (2006) with the title "A language assessment analysis of English national examination in Indonesia" is about the suitability of the English National Examination test with the language assessment principle. The participants here are students in class 12 senior high school. The method used is descriptive qualitative approach because it revealed the discussion into the descriptive one. The finding shows that English National Examination in Indonesia is not matched with some language assessment principles because there are some weaknesses in some principles.

The previous research by Grumilah & Aji (2016) focuses on the analysis of someone's perspective toward a type of assessment and the research by Sabrina (2016) analyzes using language assessment principle. Actually, this study is close to the research by Paramartha (2017) because the research focuses on investigating the quality of the multiple-choice test. However, the previous research investigates a test in offline test environment. This research investigates

the quality of an online test at junior high school level. Moreover, the previous research used jMetrik analysis while this research used an item indices analysis.