

# **CHAPTER I**

## **INTRODUCTION**

In this first chapter, the researcher wants to provide several sub-chapters. There are five sub-chapters, namely the background of the research which discusses the reasons why this research was conducted, the formulation of the research problem, the research objectives, the significance of the research, and the outline of the research. Further discussion of the sub-chapters will be written below.

### **A. Background of the Study**

In teaching and learning activities, evaluation is an important aspect. In the field of education, evaluation has a very important role. The information obtained through the evaluation will be very useful for making improvements in the future. In the formal education system, the teacher is one of the figures responsible for the learning process, whether successful or not. A good teacher not only knows how to teach but the teacher must know how to evaluate as well as how to teach. In the teaching process, a teacher must evaluate the progress of students on mastery of the lessons that have been taught within a certain period of time. The results of the evaluation will provide information about the quality of teachers and students' abilities.

The term evaluation is described as the process of collecting data on a regular basis and then interpreting or organizing it in such a way that the corresponding data can be used to establish how beneficial the teaching or system is, as well as the extent to which it is accomplishing its stated goals and expected

outcomes (Yambi, 2020, p. 6). The value of the evaluation, Brown (2010) state that should be displayed to the students in the such a way that the worth of the achievement is delivered to the students generally with several reference to the performance's consequences good or bad. Because of that, it is critical to assess students' English proficiency through an evaluation, particularly a test. Allah SWT said in the Holy Quran (Al-Mulk: 2):

الْغَفُورُ الْعَزِيزُ وَهُوَ عَمَلًا أَحْسَنُ أَيُّكُمْ لِيَبْلُوكُمْ وَالْحَيَاةَ الْمَوْتَ خَلَقَ الَّذِي

From the verse above, the creation of humans and their lives is also part of the test created by Allah for humans to be able to see his grace and power, also the creation of problems faced by all humans, Allah created for humans to evaluate themselves so that they can continue to be better in the future. In terms of education, this must be done through evaluation, one of which is by distinguishing which items are valid and reliable and which are not.

According to Putri, Suparman, & Ramlan (2015) a test is offered as means of evaluation to obtain information about the student's progress and to evaluate the outcome of the educational process (p. 2). So, evaluation based on Srinivasan (2016) suggests a layer of value judgment to the assessment process (p. 2). It is involved with the implementation of its observations and entails some assessment of a product's efficiency, utility value, or suitability in terms of cautiously structured and approved values or objectives. According to Astuti (2020), the items meet three main criteria (p. 100-101): validity, reliability, and practicality, a test

can be considered good or reasonable. To begin with, validity denotes that the test is accepted by the government. Second, measurement stability is referred to reliability. The last, practicability requires that the test be efficient.

Evaluation in education can be assumed as a formal and informal of examining students' achievement. Informal evaluation usually occurs by the time of teaching and learning process taking place. Teachers can evaluate the students' achievement by observing and making judgment based on students' performance during the process of teaching and learning. Yet, teachers cannot assume that students who never perform actively during the teaching and learning process do not understand the materials at all. It is because somehow students do not feel free to express their ideas. In terms of academic skills why students do not feel free because some reason such as based on Coplan & Evans (2009) it has been argued that shy student's academic performance may be harmed by their lack of involvement, particular use of language, performance anxiety, or reduced academic engagement (p. 211-215). Indeed, timid students tend to speak less in class, and when they do, they offer fewer comments and take longer to answer. The teacher's passion and temperament, on the other hand, have a direct influence on the students in the classroom. If the teachers wish to create a cheerful environment for their students, they should be cheerful when entering the classroom. When students perceive this level of teacher contentment and pleasure, they quickly feel joyful just simply witnessing their teacher's delight (Pekrun, 2000, pp. 143-163). Thus, it needs a formal assessment to examine the students' understanding. In contexts of education, this issue should be addressed through an

assessment which one of the goals is to distinguish between students who learn the material and those who do not.

In this study, summative items are chosen as a sort of examination that is administered at the end of a unit, semester, or year of study to assess what has been accomplished both individually and in groups. The main purpose of the summative test according to Sudijono (2008) is to determine the value that represents the success of students after going through the learning process within a certain period of time, and can determine (p. 72); the position of students in the group, students can or cannot follow the next teaching and learning process and students' progress.

However, analyzing items examination is a sequence of actions that must be completed. It is essential for teachers or team constructors to recognize which items exams are defective and ineffective. Teachers or team constructors must enhance item quality by examining the essential components of each item.

Based on Hughes (2003) state that a test's validity, in order to be considered valid, the test must first be reliable as a measuring instrument (p. 11). Valid findings need reliability, although reliability can be acquired without validity.

Based on the description above, the researcher interested in conducting research in analyzing the English summative test items for Ninth Grader of junior high school at Hulu Sungai Utara regency on academic year of 2020/2021, because the validity and reliability of the items at the school have never been tested for validity and reliability, even though the items to be tested are substitutes for national exam items that should be tested according to national standards before being given to students. This study focuses on SMPN 3 Sungai Pandan because when the

researchers conducted the research, the situation during the pandemic in all regions is getting worse so that the government implements PPKM (Enforcement of Community Activity Restrictions) level 4 which requires the implementation of teaching and learning activities in education units to be carried out through distance learning in accordance with government regulations. As a result, almost all schools closed access to this research. Only SMPN 3 Sungai Pandan gave access or permission for this research to be carried out.

So that, the researcher wants to write this thesis proposal entitle “Reliability and Validity on English Summative Test Items of Junior High School”. Therefore, this title means analyzing the summative test items of English subject of Junior High School, in order to find out the validity and reliability of items’ construction to be revised for the quality development of the test.

The reason for choosing this title is based on three main criteria. First, analyzing the test is an activity that can help improve the quality of the test. Second, it is important for teachers to design good tests. A test can be said to be a good test if it meets some of the requirements of a good test. If it does not meet the requirements of a good test, the teacher is better off or has to redesign and rearrange the test to make it better in the future. Therefore, we need to measure the quality of the test. The last, because compiling good summative test items is more difficult and more time consuming than formative tests. Summative tests must measure students’ ability to the material that has given.

**B. Research Questions**

1. How is the validity of English summative test items at SMPN 3 Sungai Pandan on academic year of 2020/2021?
2. How is the reliability of English summative test items at SMPN 3 Sungai Pandan on academic year of 2020/2021?

**C. Objectives of Study**

The purpose of this study are:

1. To identify the validity of English summative test items at SMPN 3 Sungai Pandan on academic year of 2020/2021.
2. To identify the reliability of English summative test items at SMPN 3 Sungai Pandan on academic year of 2020/2021.

**D. Significance of Study**

1. Theoretical:

The findings of this research provide a significant information about the validity and reliability of summative test item of English subject at SMPN 3 Sungai Pandan on Academic Year of 2020/2021. It is expected to be an input to improve the quality of summative test item of English subject. Furthermore, this research may be able to provide a significant contribution to other researchers as a reference for future studies on a comparable issue.

## 2. Practical:

### a. For the teacher

The results of this study can be used by teachers as data reviewers to evaluate students' determination and ability by using summative test item of English subject and as information to teachers who make items in an effort to improve the quality of the test.

### b. For the researcher

Hopefully, this research would be useful as an input to other study dealing with the scope of summative test item of English subject. Moreover, it is also arranged to finish the study at English Education Department as the final project. Furthermore, as input for researchers themselves to increase knowledge about educational evaluation.

### c. For the other researcher

This research is also expected to deliver various benefits for researchers who are going to finish their studies. And also, it is written as a consideration for further researchers who want to conduct research on the same object but different points of view.

## **E. Definition of Key Terms**

1. Item Analysis: Item analysis procedure conducted by the researcher. The process of collecting information about the quality of each test item to be tested and as an effort to find information about the validity and reliability of school final

examination item for English course. In other words, researcher will analyze the validity and reliability of school final examination item of English courses.

2. Summative Test: The activity carried out by the education unit to acquire acknowledgment of learning accomplishment and is one of the qualifications for graduation from the education unit as a substitute for the National Examination (UN).

3. English Subject: One of the junior high school subjects and one of the obligatory subjects