

## CHAPTER I

### INTRODUCTION

#### A. The Background of Study

Nowadays, evaluation takes an important role in education. Evaluation can measure the teacher' and students' progress in learning teaching process. The aim of evaluation itself is to evaluate students' achievement and students' progress in teaching and learning process. Testing is one of ways to evaluate students' ability. Teaching learning process was successful or not can be seen from the result students' test.

Testing and Evaluation of language skills and competencies are very important components of language teaching. According to Chen Desheng and Ashita Varghese (2013:31), testing becomes an integral part of teaching because it provides significant information or inputs about the growth and achievement of learner difficulties styles of learning anxiety levels. It has an important role to measure the students' achievement during teaching and learning process. Thus, the test maker or test constructor should be able to construct a good test. Teachers who construct a good test will give a good contribution to student's education. On the other hand, teachers who have lack of skill in constructing a good test will give less contribution or might even make student's education become worst.

There are some items of the test; they are multiple choice, essay, or even oral test, but the most common form of items used by the teacher for junior high school and senior high school are multiple choice and essay. The first form is

multiple choice tests. This kind of test item is considered to be simpler than essay in the form of scoring method because teachers only count the correct answer. The second form is essay test. By conducting an essay test, teachers will know how deep the students understand on the material because in the form of essay test, students will not be provided by the optional answers but they must create their own answer by their own language. So, from the students' answer, teachers will be able to know whether the students are really understood. However, good multiple-choice test and essay test are not easy to make because the test designer must consider about its validity, reliability, index of difficulty and the last is the discrimination item. Thus, it is very necessary to create a good test items for students because the result of the test will be used as the representative of the students' ability so that the test items must be valid and reliable and it must also comply with the characteristics of a good test in term of the level of difficulty, and discrimination item.

Dr. Foyewa, R.A (2015:32) stated that there are certain qualities expected of a good test. They include among others :

a. Validity : A good language test should measure what it supposed to measure. There are different types of validity. These are; face validity, contents validity, predictive validity, concurrent validity and constructed validity.

1) Face validity refers to how accurately a conclusion, measurement, or concept corresponds to what is being tested. The example of the test should has a good face validity is objective test such as multiple choice test and make a match test.

- 2) Contents validity is refers to how well a test measures the behavior for which it is intended. Example of contents validity is psychology test.
  - 3) Predictive validity is the extent to which a score on a scale or test predicts scores on some criterion measure, example: SBMPTN test should able to predict the ability of students their ability in particular major in university.
  - 4) Concurrent validity It refers to the extent to which the results of a particular test, or measurement, correspond to those of a previously established measurement for the same construct. Example of test which have a good concurrent validity is National Examination test.
  - 5) Construct validity is the degree to which inferences can legitimately be made from the operationalizations in the study to the theoretical constructs on which those operationalizations were based. Example of construct validity is Summative test and Final test.
- b. Reliability: Reliability is the quality of being reliable: language test reliability is the consistency of a test in measuring what it is supposed to measure. A good language test is expected to be reliable.
  - c. Objectivity: this quality of a language test ensures that a test should have one and only one correct answer. Examples of this include the “multiple choice” and “true and false” tests.
  - d. Economy: this quality of test ensures that the cost of administering a test, the time involved in setting and marking it should be commensurate with the expected result obtained from it. A test that takes much time, energy and costs much to construct cannot be said to be economical.

A problem arises when most teachers underestimate an evaluation of the test item in the English final test they have made, whereas, this evaluation is in fact so important for the teachers in order to know the quality of the test they made; whether it is already valid and fulfill the characteristics of good test or not. Analyzing test items include analyzing the validity, reliability, level of difficulty, discrimination power.

There are two categories to make test. They are standardized test and teacher made test. Standardized test is a test which is made by professional testing service that the test tries to analyze and revise before being tested. Commonly a teacher made test has a lower reliability and validity than standardized test because in making a test there is no much consideration that teacher made, thus in need more to be analyzed. Before conducting her research the writer did her preliminary research in the location of her research location. The preliminary research was conducted on February 2017. In this case, the researcher concerned with the analysis of English teacher's made mid test for second grade in academic year 2016/2017. The mid test consisted of 40 items. After analyzing the test, it was found that the test didn't have a good validity because from those 40 items there are 17 items (42.5 %) that do not fulfill the requirements of validity. The example of item invalid is number 2, after analyzed it, the t table is less than r table. The t table is 0,334 taken from product moment while the r table found only 0,156. Then, considering the reliability, the test also found still not reliable, it shown from the data that the value obtained is  $0,668 < 0,70$ , so the test item of the second grade students of SMA N 1 Pagaran in the academic year of 2016/2017 is unreliable. Then, in measuring the index of difficulty of test, the researcher found

that there are 30 item out of 40 item are easy, because they have of difficulty value between 0,71-1,00. Those items need to be revised because it is too easy for students to answer it moreover it is not belong to the criteria of the good test. The last thing that need to be considered is the discrimination power of the test. The result's analysis showed that there are 14 items (35 %) has a poor index of discrimination and those items must be revised. There are 15 (37,5 %) has satisfactory. Those items already accepted in the test. in addition, there are 1 item (2,5 %) has a good index of discrimination, and there are 10 items (25 %) has wrong index discrimination. Those ten items are rejected and cannot be used in the test.

Based on the preliminary research and based on writer's observation and interview with the English teacher in SMA N 1 Pagaran. The writer choose SMA N 1 Pagaran as the location of her research because the test made by the group of English teacher need to be revised. They mostly take the test from internet, LKS, previous exercise and sometime from previous test. In addition in constructing the test there is no much considered they made, such as whether the test can be as the representative of the students' ability during almost 3 years gaining the knowledge in the school. The writer can take a conclusion that the quality of the test made by the group English teacher at SMA N 1 Pagaran particularly the final test intended to third grade students is still unknown, thus the writer wants to analyze the quality in order to know whether the test deserve to be tested or not. In addition, by analyzing those aspects, teachers will know whether the test they made is already valid and reliable or not, whether the test is too easy or too difficult for the students, whether the test can discriminate the upper and lower students or not and

teachers should make sure that the distractors they made are really able to distract students' answer because the better the test items constructed by the teacher is the more reliable the score of the students and the reliable score can be used as the representative of the students' ability.

Considering the explanation above, this study focus on analyzing the quality of English final test at third grade of students at SMAN 1 Pagaran concerning study about the validity, reliability, item analysis includes index of difficulty, and index of discrimination. The test is final school examination which constructed by the group of English teacher in SMA N 1 Pagaran. The form of the test is multiple choice form. The test can be seen in appendix A. To analyze those aspects the writer could do either manually or using application of computer. Manually means to measure each of aspects the writer use formula stated or to make it easier and faster the writer can use the application of SPSS. In this study, the writer used SPSS version 17.

### **B. Problem of the Study**

Based on the background of the study, the problem of this study is: how is the quality of English final test made by the teacher for second semester of third grade students at SMA N 1 Pagaran in academic year 2016/2017 ?

### **C. The Objectives of Study**

The objective of Study is to find out the quality of English final test made by the teacher for second semester of third grade students at SMA N 1 Pagaran in academic year 2016/2017 ?

#### **D. The Scope of Study**

The scope and limitation of this study is English final test for third grade at SMAN 1 Pagaran. The test was made by group of English teacher in SMAN 1 Pagaran. The test consists of thirty five multiple choice questions. In this study the researcher will focus on the quality of teacher made English Final test items based on characteristic a good test include validity, reliability, and item analysis (index difficulty, and index discrimination).

#### **E. Significances of the Study**

Related to the objectives of the study, this analysis is intended to see some advantages as elaborated in some paragraphs below. There are three major significances that this study wants to contribute.

The first one is theoretical significance. This study may give basic understanding to the teachers, educators, trainers, and others that assessment and evaluation cannot be made and assumed only by basing on students or one's outer performance or guessing in some cases. They should know that the test items should be made to evaluate students' understanding and ability. The tests are also useful to develop their professionalism as being an educator.

The second one is practical significances. This study is beneficial for the test makers as additional reference in constructing and analyzing test items and their procedures.

The last one is pedagogical significance. This study provides English teachers with some meaningful and useful information of efficient class

discussion of the test result, the general improvement of classroom instruction, evaluation in teaching learning process, and improvement in test making.



THE  
*Character Building*  
UNIVERSITY