

Chapter Two

Literature Review

Definition of Test

Test is examination, evaluation, observation, trial or used (under actual or simulated operating conditions or environmental) to Determine and document (1) capabilities, characteristics, effectiveness, reliability, and / or suitability of a material, product, or system, or the (2) ability, aptitude, behavior, skill level, knowledge, or performance of an person.

There are some definitions about test from some expert. Ur (1996) said that tests are used as a means to motivate students to learn or review specific material. It means test is one motivation of students to learn or review material in their school. Furthermore Fernandes (1984) states that a test is a systematic procedure for surveying a person's behavior and explaining it with the aid of a numerical scale or a category system. While Airisian & Russel, (2008) says that test is a formal, systematic procedure used to gather information about students achievement or other cognitive skill. In addition, Linn & Gronlund (1995) argued that a test is a particular type of assessment that typically consists of a set of questions administered during a fixed period of time under reasonably comparable conditions for all students.

In summary, a test is the particular types of assessment to reinforce learning and to motivate students by giving a task or a set of tasks. Through the test, the lecturer do not only measure and motivate the students' ability but also

improve the lesson in teaching learning process. In order to make a proper decision, the lecturer needs accurate data and to gain data, so a good instrument is needed.

Language testing

Language testing is a device that tries to assess how much has been learned in a foreign language course or some part of course (Oller, 1979). For any students whose native, language or language variety is not used in the schools, many tests are not traditionally taught as language test, it may be primarily tests of language ability.

According to Mcnamara (2000) language testing is a procedure for gathering evidence of general or specific language abilities from performance on tasks designed to provide a basis for predictions about an individual's use of those abilities in real world contexts. All such test require us to make distinction between the data of the learner's behavior, the actual language that is produced in test performance, and what this data signify, that is to say what they count as in term of evidence of proficiency, readiness for communicative roles in the real world.

The same basic understanding of the facts of language learning applies to language testing. What the student has to learn constitutes the corpus of what we have to test. Since the students have to learn language, it is language that we must test (Lado, 1961).

There are several types of test for measuring the skills, for example, placement test, achievement test, proficiency test and aptitude test. Each test has

its functions depend on the institution or companies that apply it. In this research, the test is use to be analyzed is progress test. It will start from the definition of the progress test. Progress test is test to check the students' progress in learning process (Arthur, 2003). The content of tests based on the course objectives gives a number of advantages.

The first, it compels course designers to be explicit about objectives. The second, it makes possible for performance on the test to show just how far students have achieved the instructional objectives. Consequently, the course designer or teacher should construct a syllabus based on the instructional objectives and should select books and materials which are consistent with the course objectives.

Progress test is a test made for measuring how far student masters material taught in the classroom. It is often given to motivate students. They also enable students to assess the degree of success of teaching and learning and to identify areas of weakness & difficulty.

Progress test is designed by two types: short term and long term. The short term progress test is to check how well the students understood or learned material covered in specific unit or chapter. Thus, it enable the lecturer to make remedial or not if it is required. While, the long term progress test is to check the learners' progress over the entire course. It enables to judge how well the students have progressed. Administratively, it uses for sole basis of decision to promote a higher level.

Criteria of good test

There are many considerations entering into the evaluation of a test, which referred as a good test because a good test can provide available information for a good evaluation in order to measure student's comprehension of instructional objectives, but the writer consider them under three main heading.

These are respectively validity, reliability, and practicality. Validity refers to the extent to which a test measures what we actually wish to measure.

According to Brown (2001, pp. 386-387) Validity test actually measures what is intended to measure. Reliability is constant and reliable and practicality is meant as financial limitations, time constraints, ease of administration, and scoring and interpretation.

In other hands, the researcher says that a good test should have three components and its supporting parts in order to test consist of good standard.

Validity

The single most important characteristic of a good test is its ability to help the lecturer make a correct decision of what is intended to measure. This characteristic is called validity. Validity is concerned with whether the information being gathered is relevant to the decision that needs to be made.

A test has validity if it measures appropriately, what it is supposed to measure. According to Heaton (1977, p. 159) the validity measures the test to measure and nothing else. Finnachiaro & Sako (1993, p. 24) also stated that a test

measures effectively what it is intended to measure so it is valid. Still in the same sense, Tinambunan (1998, p. 11) states that the validity is the extent to which the test measures what is intended to measure. Also, Gronlund (1968, p. 105) states that test scores are valid to the extent to which they serve the use for which they are intended. While J. Staley Ahman & Glock (1967, p. 285) point out that in educational measurement, validity is often defined as the degree to which a measuring actually serves the purpose for it is intended.

In reviews, validity of test is important to know whether a test has a good quality in testing someone's capacity. As the validity is one of the most important characteristic of test scores, the constructor of the test should know the various aspects from the validity itself and various procedures by which they are determined.

According to Gronlund (1968) the two most important characteristics of test scores are validity and reliability. Anyone working with tests-whether constructing them or using published tests-should understand the meaning of these concepts and should know the various procedures by which they are determined.

According to Heaton (1977, p. 159), a validity of a test can be seen from some aspects mentioned below.

Content Validity

A test has content validity if the test contains materials that the students have been taught. To fulfill this, the lecturer also should refer to the instructional objectives of teaching learning process. Finnachiaro & Sako (1993) stated that

content validity is solid by checking all items in the test to make certain that they correspond to the instructional objective of the course. Still in the same sense, Noll (1965, p. 79) explains that the test has content validity if a lecturer gives a test which deals with the material and with the objectives of instruction in particular.

The researcher concludes that content validity measures evidence that the items are comprehensive related to its intended measurement concept.

Construct Validity

A test is said to have construct validity if it can demonstrate that it measures just the ability, which it is supposed to measure. If test has construct validity, it is capable of measuring certain specific characteristic in accordance with theory of language behavior and learning.

In short, the study of how the test results are influenced by changes in individuals. It should respectively influence or fail to influence the individuals' position.

Concurrent Validity

Concurrent validity can be measured by comparing the extent to which test scores correlate with scores on other relevant measures administered at the same time.

Concurrent Validity needs to be examined whenever one measure is substituted for another, such as allowing students to pass a test instead of taking a

course. Concurrent validity is determined when test scores and criterion measurement are either made at the same time (concurrently).

Face validity

A test has face validity if the test has a good face or the way the test looks, if a test item looks right to other testers, lecturers, moderators, and students, it can be described as having at least face validity. While Finnachiaro & Sako (1993) defined that test judgments based on the way the test looks to educators, students, and general public. The test should not only 'be right' but also 'looks right'.

In review, the researcher can say that face validity measures appearance to be assessing the intended construct. It may be an essential component in enlisting the motivation.

However, due to some limitations that the researcher has undergone, the researcher then only chose two types of validity to analyze; content validity and concurrent validity.

The Importance of Test Validity

Test is usually used for measuring the students' skill or abilities. It measures how far the students have mastered the materials that have given during the course. It is also an important aspect to know is there any impacts to individual person or not. Regardless of the reasons, test should have a good quality because the test perhaps appropriate in one situation but it is inappropriate in another. One of the criteria of a good test is that test must be valid. To check

whether or not the test is valid, the researcher analyzed the test items by considering the contents, concurrent, construct and face validity.

Validity is the extent to which a measurement tool measures what it supposed to measure. The several types of validity should be used to build a case for valid test use. It could be done by doing the comparing the test specifications with the course requirement to see if there is sufficient overlap to be comfortable using evidence. It is to complete the concurrent criterion to determine the relationship between course grade and test score and also follow up with surveys of the students enrolled to determine whether they felt their preparation to be adequate.