**Assessment and Testing**

**1. Assessment, testing, Measurement, and Evaluation:**

Many think that assessment and testing are synonyms, but they are not. A test is an instrument or procedure designed to elicit performance from learners with the purpose of measuring their attainment of specified criteria. Tests are a subset of assessment, a genre of assessment techniques They are almost always identifiable time periods in a curriculum when learners muster all their faculties to offer peak performance, knowing that their responses are being measured and evaluated.

Douglass Brown defines a test as “a method of measuring a person’s ability, knowledge or performance in a given domain.” A test is intended to measure knowledge and skills one acquires during a sequence of sessions. (it is an instrument that’s used to measure student’s learning at a particular point in time). In short, it is an assessment instrument that is chiefly intended to turn students’ performance into numbers. (e.g. multiple-choice tests, quizzes, cloze tests).

Assessment, on the other hand, is a more encompassing term than testing. It is the process of gathering, interpreting, and sometimes recording and using information about students' responses to an educational task in order to provide the next learning step.

Assessment is an ongoing process that includes a wide range of techniques. Whenever a student responds to a question, offers a comment, or tries out a new word or structure, the teacher subconsciously makes an appraisal of the student’s performance. Written work-from jotted down phrase to a formal essay-is performance that is ultimately ‘judged by self, teacher, and possibly other students. A good teacher never ceases to assess students, whether those assessments are incidental or intended.

In the most general sense, assessment is the process of making a judgment or measurement of worth of an entity (e.g., person, process, or program). Educational assessment involves gathering and evaluating data evolving from planned learning activities or programs.

It is the process whereby teachers gather information about students’ performance or  
ability.

Measurement is the process of determining the amount or length of something when compared with a fixed unit (e.g. using a ruler to measure length). In language teaching, measurement constitutes the quantification of language proficiency. It is the process of quantifying the observed performance of learners

Yet another potentially confusing term that needs explanation is evaluation. Is evaluation the same as testing? Evaluation does not necessarily entail testing; rather, evaluation is involved when the results of a test (or other assessment procedure) are used for decision making (Bachman, 1990, pp. 22-23). Evaluation involves the interpretation of information. Simply recording numbers or making check marks on a chart does not constitute evaluation. You evaluate when you "value" the results in such a way that the worth of the performance is conveyed to the test-taker, usually with some reference to the consequences-good or bad-of the performance. Test scores are an example of measurement, and conveying the "meaning" of those scores is evaluation. If a student achieves a score of 75 percent (measure-ment) on a final classroom examination, he or she may be told that the score resulted in a failure (evaluation) to pass the course.

Evaluation may focus on the effectiveness or impact of a program of instruction, examination or project. Students are usually not asked to evaluate while teachers carry out or take part in evaluation only in some contexts. ‘Experts’ or the authorities are most commonly legitimized to carry out formal evaluation.

**2. Assessment Types:**

It is important to distinguish between two forms of assessment, formal and informal.

**Informal assessment** is involved in all incidental, unplanned evaluative feedback on tasks designed to elicit performance, but *not* for the purpose of recording results and making fixed judgments about a student's competence. Informal assessment is what testing experts call formative evaluation. Formative assessment often implies the observation of the process of learning, as opposed to the product. Teachers’ success is greatly dependent on constant informal assessment for it gives learners information about how they are progressing toward goals and what the next step in the learning process might be. It is assessment for teaching and learning.

**Formal assessment**, also called summative assessment, refers to making judgments about the performance or ability of students by assigning grades to them. Summative assessment occurs periodically; either at the end of year, semester, or two or more units. Representatives of summative assessment can be quizzes, exams, tests, and end-term papers. The focus of summative assessment is directed towards the final product rather than the process. It is assessment of teaching and learning.

**3. Purposes of Assessment:**

**Information collected in assessment help  
1. teachers:**

- to understand how students are progressing;  
- recognize students’ strengths and areas for improvement in learning;  
- work out ways of helping students;  
- plan their lessons

**2. students to**  
- understand their own progress;  
- recognize areas and ways for improvements in learning

**4. Types of Tests:**

**1. Proficiency test**:

A proficiency test is devised to measure how much of a language someone has learned. It is a type of test that measures the learners’ general level of language mastery. Some proficiency tests have been standardized for worldwide use, such as the well-known American tests, the TOEFL, and the English Language Proficiency Test (ELPT) which are used to measure the English language proficiency of foreign students intending further study at English-speaking institutions, namely the USA. As Valette (1977) posits, “*the aim of a proficiency test is to determine whether this language ability corresponds to specific language requirements”,*getting a job or studying abroad.

**2.** **Placement test:**

A test that is designed by institutions for new students. It aims at putting test takers into groups so that they roughly have the same level as other students in the group when they start studying**.** It places students based on their level or grades.

**3.** **Diagnostic test:**

Diagnostic tests are used to pinpoint the strengths and weaknesses of the students. They are usually designed to guide remedial instruction as they seek to identify those  
areas in which a student needs further help, and they take place at the beginning of the year, course, and unit.

**4. Achievement tests:**

It measures what has been learned in a course or a sequence of courses; it measures the extent to which test takers learned the material presented in a course. This type of test means to assess what students have learned after completing a unit.

**5. Aptitude test:**

Itis any type of assessment that evaluates the talent, the ability, potential  
to perform a certain task, in other words, to discover whether a student has talent or  
basic ability for learning new language or not.

**5. Criteria for designing and selecting an effective test**

In order for a test to be effective, it should have the following characteristics:

**1. Authenticity:**

Basically, using the principle of authenticity in a test uses the test for performing an assignment in a real life situation. Thus, in a test, authenticity may be presented in the following ways; a) includes as much natural language as possible b) contains components which are contextualized c) has meaningful, relevant, real life themes d) provides some thematic organization to items, such as a story line or episode e) offers assignments that replicate real world assignments (Brown, 2010).

**2. Reliability:**

A reliable test is a test that is characterized by consistency. It is a test that renders the same (consistent) results when it is administered to the same students in two occasions. (the same environment, timing, no external factors intervened in that test, should yield the same results).

Reliability is most often estimated with regard to:

* **Internal reliability:** it means the degree of the interrelationship among the items of a test. It reflects the extent to which items within an instrument measure various aspects of the same characteristic or construct.

Example: Designing a test about geometry, then all questions should be about geometry.

* **Inter-rater reliability**: it refers to the level of agreement between two or more evaluators/ judges/ raters on a particular instrument at a particular time.

**3. Validity:**“Validity of assessment is perceived when the assessment method is consistent with the material and curriculum being taught and if the results of the assessment are accurate” (Brown, 2002; Gür, 2013). Validity is an important feature in designing an effective test; it refers to the extent to which a test measures what it intends to measure. The saying ***“Test, what you’ve taught”***, illustrates what validity means.

**Content validity:** A test is said to have content validity if the items or tasks of which it is made up constitute a representative sample of items or tasks for the area of knowledge or ability to be tested (often related to a syllabus or a course).

**Face validity:** it is about whether a test appears to measure what it is supposed to measure.

4. **Practicality:**A practical test is a test that usually does not take too much time for the students to accomplish, has a specific and time-efficient evaluation procedure, that is, just as the test should be conducted within a timing that is appropriate (5 hours for a test is too much), the test should be relatively easy to correct for the teacher, a test that takes a teacher too much time to correct is impractical. The instructions should be clear and marks allotted to each items should be reasonable.

**The principles of practicality are encapsulated in the four statements:**  
⮚Is not excessively expansive.  
⮚Stays within appropriate time constraints.  
⮚Is relatively easy to administer.  
⮚Has a scoring/evaluation procedure that is specific and time-efficient.

**5. Washback effect:**Washback effect refers to the influence a test has on individuals. More specifically, it refers to  
the influence the test has on the test takers, students, and the test designers, teachers. Students,  
for instance, would realize the mistakes they have made in a test after its correction and rectify them. The same applies to teachers who may as well work on improving or changing their pedagogical strategies.

To sum up, the significance of assessment in the process of language teaching has been known recently. Assessment is very important for the students to acquire a language. It plays a crucial role in the process of learning, and connects students to new knowledge using their current abilities.