

Tikrit University
College of Education for Humanities
English Department



PhD (Second Course) 2025-2026

**An Advanced Course in EFL Communicative Testing, Measurement and
Evaluation**

Assessment Concepts and Issues(Part One)

(1)

Prof. Dunia Tahir (PhD)

Brown, H. D., & Abeywickrama, P. (2018). *Language Assessment Principles and Classroom Practices* (3rd ed.). Pearson Education.

2026.A.D.

1447.A.H.

Chapter 1

Assessment Concepts and Issues

Assessment and Testing

Assessment is a popular and sometimes misunderstood term in current educational practice. **Assessment** is “appraising or estimating the level or magnitude of some attribute of a person”. **In educational practice**, assessment is an Ongoing process that encompasses a wide range of methodological techniques. Whenever a student responds to a question, offers a comment, or tries a new word or structure, the teacher subconsciously appraises the student’s performance.(Brown,2003,p.4)

Assessment is the process of collecting information about learners using different methods or tools . e.g. tests, quizzes.

Educators assess their students for a variety of purposes:

a-to evaluate learners' educational need.

b- to diagnose students' academic readiness.

c-to measure their progress in a course.

d-to measure skill acquisition.

Q1:What is assessment?

Q2:Is assessment the same as testing?

Q3:How is assessment used in educational practice?

Q4:When does teacher assess students?

On the other hand, **A test** is one form of assessment and refers to procedures used to measure a learners' learning at a specific point in time and often involves collective information in numerical form. Common forms of tests are multiple choice questions and gap-fill or cloze tests. In English classes, teachers also need to assess their students' learning to determine the effectiveness of their teaching and of the materials they use.(ibid)

Tests are a subset of assessment, a genre of assessment techniques. They are prepared administrative procedures that occur at identifiable times in a curriculum when

learners muster all their faculties to offer peak performance, knowing that their responses are being measured and evaluated.

In scientific terms, **a test** is a method of measuring a person's ability, knowledge or performance in a given domain. The components of this definition are: A test is

first a method. It is an instrument, a set of techniques, procedures or items that requires performance on the part of the test-taker. To qualify as a test, the method must be explicit and structured: multiple choice questions with prescribed correct answers, a writing prompt with a scoring rubric, an oral interview based on a question script, or a checklist of expected responses to be filled in by the administrator.

Second, a test must **measure**, which may be defined as a process of quantifying a test-taker's performance according to explicit procedures or rules.

Some tests measure general ability, where as others focus on very specific competencies or objectives. A multi skill proficiency test determines a general ability level; a quiz on recognizing correct use of definite articles measures specific knowledge.

Third, a test measures **an individual's** ability, knowledge, or performance. Teachers need to understand who the test-taker are. What are their previous experiences and backgrounds? Is the test appropriately matched to their abilities?

Fourth, A test measures **performance**, but the results imply the test-taker's ability or to use a term common in the field of linguistics, **competence**. Most language tests measure one's ability to perform language, that is, to speak, write, read, or listen to a subset of language. tests are occasionally designed to tap into a test-taker's knowledge about language: Performance-based tests sample the test-taker's actual use of language, and from those samples the test administrator infers general competence.

Finally, a test measures **a given domain**. For example in the case of proficiency test , even though the actual performance on the test involves only a sampling of skills, the domain is overall proficiency in a language, general competence in all skills of a language. Other tests may have more specific criteria. A test of pronunciation might well be a test of only a limited set of phonemic minimal pairs. A vocabulary test may focus on only the set of words covered in a particular lesson or unit.

A well-constructed test is an instrument that provides an accurate measure of the test-taker's ability within a particular domain.

Q5:What are tests ?

Q6:In scientific terms, what is a test?

Q7:Why is a test considered a method?

Q8:What types of test methods are mentioned in the text?

Q9:What must a test do besides being a method?

Q10:What is the difference between general ability and specific knowledge tests ?

Q11:What does a test measure?

Q12:Why must test-takers be understood by test designers?

Q13:What does a test measure: performance or competence?

Q14:What is competence in linguistics?

Q15:How do performance-based tests work?.

Q16: Give an example of a performance-based language test..

Q17:What is meant by measuring a given domain?

Q18:Can a test measure all language skills at once?

Q19:What is one major challenge in test construction?.

Q20:What is a well-constructed test?

Q21:Is constructing a good test simple?

Measurement and Evaluation

According to Brown (2003), Measurement is the process of quantifying the observed performance of classroom learners. Bachman (1990) cautioned us to distinguish between quantitative and qualitative descriptions of student performance. Simply put, the former involves assigning numbers (including rankings and letter grades) to observed performance, whereas the latter consists of written descriptions, oral feedback, and other nonquantifiable reports.

Quantification has clear advantages. Numbers allow us to provide exact descriptions of student performance and to compare one student with another more easily.

They also can spur us to be explicit in our specifications for scoring student responses, thereby leading to greater objectivity. On the other hand, quantifying student performance can work against the teacher or tester, perhaps masking nuances of performance or giving an air of certainty when scoring rubrics may actually be quite vague. Verbal or qualitative descriptions may offer an opportunity for a teacher to individualize feedback for a student.

on the other hand, evaluation is qualitative descriptions of student performance .(ibid)

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 to make decisions. 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 you convey the worth of the performance 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% (measurement) on a final classroom examination, he or she may be told that the score resulted in a failure (evaluation) to pass the course. Evaluation can take place without measurement, as in, for example, a teacher’s appraisal of a student’s correct oral response with words like “excellent insight, Ahmed”. .(Brown,2003)

Q12:What are measurement and evaluation?

Q13:What is measurement?

Q1:What is the difference between quantitative and qualitative description?

Q2:What are the advantages of quantification?

Q3:What are the disadvantages of quantification?

Q4:What advantages do qualitative descriptions offer?

Q5:Is evaluation the same as testing?

Q6:What is evaluation?

Q7:Does recording numbers count as evaluation?

Q8:When does evaluation occur?

Q9: How are measurement and evaluation related?

Q10:Can evaluation occur without measurement?

Q11:Are tests the only form of assessment?

Assessment and Learning

Returning to our contrast between tests and assessment, we find that tests are a subset of assessment, but they are certainly not the only form of assessment that a teacher can apply. Although tests can be useful devices, they are only one among many procedures and tasks that teachers can ultimately use to assess (and measure) students.

But now, you might be thinking, if you make assessments every time you teach something in the classroom, does all teaching involve assessment? Are teachers constantly assessing students, with no assessment-free interactions? .(Brown,2003)

The answers depend on your perspective. For optimal learning to take place, students in the classroom must have the freedom to experiment, to try out their own hypotheses about language without feeling their overall competence is judged. learners have should have opportunities to “play” with language in a classroom without being formally graded. the opportunities for learners to listen, think, take risks, set goals, At the same time, during these practice activities, teachers are indeed observing students’ performance, possibly taking measurements, offering qualitative feedback, and suggesting strategies.

Q1:Does all teaching involve assessment?

Q2:Why should students not always be formally assessed?

Q3:How is language learning compared to sports in the text?

Q4:What do teachers do during practice activities?

Q5:What kinds of questions do teachers ask when observing performance?

Q6:How are testing, assessment , and teaching related in the classroom?

Q7:What is evaluation?

Informal and Formal Assessment

Informal assessment can take a number of forms, starting with incidental, unplanned comments and responses, along with coaching and other impromptu feedback to the student. Examples include saying, “Nice job!” or “Good work!” .informal assessment is embedded in classroom tasks designed to elicit performance without recording results and making fixed conclusions about a student’s competence. Informal assessment is virtually always nonjudgmental, not making ultimate decisions about the student’s performance.

On the other hand, formal assessments are exercises or procedures specifically designed to tap into a storehouse of skills and knowledge. They are systematic, planned sampling techniques constructed to give teachers and students an appraisal of student achievement. Formal assessment is virtually always judgmental, making ultimate decisions about the student’s performance. .(Brown,2003,p.5)

Is formal assessment the same as a test?

Formal assessment and testing are related but not the same.

1. Tests as Formal Assessments:

All tests can be considered formal assessments, but not all formal assessments are tests.

2. Examples of Formal Assessments:

- Journals and Portfolios: A student's journal or collection of work can formally assess whether they have met certain course goals, but these are not typically labeled as tests.
- Observations: Tracking how often a student participates in class discussions is another example of formal assessment, yet it doesn't fit the traditional definition of a test.

3. Characteristics of Tests:

- Tests are generally time-limited, usually lasting a class period or just a few hours.
- They typically focus on a small sample of a student's behavior or knowledge.

While tests are a type of formal assessment, other methods like journals and observations also serve this purpose but are not classified as tests.

Q1:What is Informal Assessment?

Q2:Give examples of informal assessment?

Q3: Is informal assessment always planned?

Q4: Why is informal assessment important?

Q5: How does informal assessment help learning?

Q6: Are all tests formal assessment?

Q7: Are all formal assessments tests?

Q8: Explain the statement: "All tests are formal assessments, but not all formal assessments are testing".

Q9: What is informal assessment?

Q10: Give examples of informal assessment.

Q11: What is formal assessment?

Formative and Summative Assessment

Formative, assessment: evaluating students in the process of “forming” their competencies and skills with the goal of helping them to continue that growth process. The key to such formation is the delivery (by the teacher) and internalization (by the student) of appropriate feedback on performance, with an eye toward the future continuation (or formation) of learning. For all practical purposes, virtually all kinds of informal assessment are (or should be) formative. They have as their primary focus the ongoing development of the learner's language. So, when you give a student a comment or a suggestion, or call attention to an error, you offer that feedback to improve the learner's language ability.

Summative assessment aims to measure, or summarize, what a student has grasped and typically occurs at the end of a course or unit of instruction. A summation of what a student has learned implies looking back and taking stock of how well that student has accomplished objectives, but it does not necessarily point to future progress. Final exams in a course and general proficiency exams are examples of summative assessment. Summative assessment often, but not always, involves evaluation (decision making). (Brown, 2003)

Q1: What is formative assessment?

Q2: What is the main purpose of formative assessment?

Q3: Why are most informal assessments considered formative?

Q4: What is summative assessment?

Q5: Does summative assessment always involve evaluation?

Norm Referenced and Criterion-Referenced Tests

In norm-referenced tests, each test-taker's score is interpreted in relation to a mean (average score), median (middle score), standard deviation (extent of variance in scores), and/or percentile rank. The purpose of such tests is to place test takers along a mathematical continuum in rank order. Scores are usually reported back to the test-taker in the form of a numerical score (for example, 230 out of 300) and a percentile rank (such as 84 percent, which means that the test-taker's score was higher than 84 percent of the total number of test takers, but lower than 16 percent in that administration). Typical of norm-referenced tests are standardized Tests like the Scholastic Aptitude Test or the Test of English as a Foreign Language (TOEFL). Money and efficiency are primary Concerns in these tests.

Criterion-referenced tests, on the other hand, are designed to give test-takers feedback, usually in the form of grades, on specific course or lesson objectives. Classroom tests involving the students in only one class, and connected to a curriculum, are typical of criterion-referenced testing. Here, much time and effort on the part of the teacher (test administrator) are sometimes required in order to deliver useful, appropriate feedback to students, criterion-referenced testing is of more prominent interest than norm-referenced testing.

Q1: What is norm-referenced tests?

Q2: How are scores interpreted in norm-referenced tests?

Q3: What is the main purpose of norm-referenced tests?

Q4: How are results usually reported in norm-referenced tests?

Q5: Why are norm-referenced tests efficient?

Q6: What are criterion-referenced tests?

Q7: Where are criterion-referenced tests commonly used?

Q8:What is the main focus of criterion-referenced tests?

Q9:Does score distributions matter in criterion-referenced tests?

Q10:Which type of test requires more teacher effort?

Q1: What is the main purpose of assessment instruments?

Types and Purposes of Assessment

Q1:What is meant by classroom-based assessment?

It refers to assessments designed and used by teachers to measure achievement or diagnose difficulties within a lesson or course.

Q2:Why is it important to specify the purpose of an assessment?

Because defining the purpose and objectives is the first essential step in choosing, designing, revising, or adapting an assessment procedure.

Q3:Into how many types do tests generally fall?

Tests fall into a finite number of types, classified according to their purpose.

Achievement Tests

The most frequent purpose for which a classroom teacher uses a test is to measure learners' ability within a classroom lesson, a unit, or even an entire curriculum. Commonly called achievement tests, they are limited to particular material addressed in a curriculum within a specific time frame . Achievement tests can also serve the diagnostic role of indicating what a student needs to continue to work on in the future, but the primary role of an achievement test is to determine whether course objectives have been met and appropriate knowledge and skills acquired by the end of a given period of instruction.

Achievement tests are often summative because they are administered at the end of a lesson, unit, or term of study. They also play an important formative role, because an effective achievement test offers feedback about the quality of a learner's performance in the unit or course.

The specifications for an achievement test should be determined by the:

* objectives of the lesson, unit, or course being assessed

- * relative importance (or weight) assigned to each objective
- * tasks used in classroom lessons during the unit of time
- * time frame for the test itself and for returning evaluations to students
- * potential for formative feedback

Q1:What is an achievement test?

Q2:When are achievement tests usually given ? And What do achievement tests focus on?

Q3:What is the primary role of an achievement test?

Q4:Why are achievement tests considered summative?

Q5:How do achievement tests serve a formative role?

Q6:What factors determine the specifications of an achievement test?

Diagnostic Tests

The purpose of a diagnostic test is to identify aspects of a language that a student needs to develop or that a course should include. A test of pronunciation, for example, might diagnose the phonological features of English that are difficult for learners and should therefore become part of a curriculum. Such tests usually offer a checklist of features for the administrator (often the teacher) to use to pinpoint difficulties.

The distinction between a diagnostic test and a general achievement test. Achievement tests analyze the extent to which students have acquired language features that have already been taught, diagnostic tests should elicit information on what students need to work on in the future. Therefore, a diagnostic test typically offers more detailed, subcategorized information about the learner.

Q1:What is the purpose of a diagnostic test?

Q2:What kind of information do diagnostic tests provide?

Q3:How is a diagnostic test different from an achievement test?

Q4:Why do diagnostic tests offer more detailed information than achievement tests?

Q5:How can diagnostic tests help teachers?

Placement Tests

Some achievement and proficiency tests can act as placement tests, the purpose of which is to place a student into a particular level or section of a language curriculum or school. A placement test usually, but not always, includes a sampling of the material to be covered in the various courses in a Curriculum; a student's performance on the test should indicate the point at which the student will find material neither too easy nor too difficult but appropriately challenging. placement tests can provide teachers with useful information on what may or may not need to be emphasized in the weeks to come. So, a placement test takes on a formative role.

Placement tests come in many varieties assessing comprehension and production, responding through written and oral performance, using open ended and limited responses, applying selection (e.g., multiple-choice) and gap filling formats depending on the nature of a program and its needs. the ultimate objective of a placement test is to correctly place a student into a course or level and provide diagnostic information on a student's performance.

Q1:What is the main purpose of a placement test?

Q2:What does a placement test usually include?

Q3:What should a student's performance on a placement test indicate?

Q4Why should placement tests sometimes be diagnostic?

Q5:How do placement tests take on a formative role?

Q6:In what forms can placement tests be given?

Proficiency Tests

A proficiency test used to test a student's general ability with the language. It is not intended to be limited to any course, curriculum, or single skill in the language. Proficiency tests have traditionally consisted of standardized multiple-choice items on grammar, vocabulary, reading comprehension, aural comprehension, and sometimes a sample of writing. Typical examples of standardized proficiency tests are the Test of English as a Foreign Language (TOEFL). Proficiency tests are almost always summative and norm referenced. They provide results in the form of a single score.

A key issue in testing proficiency is how the constructs of language ability are specified. Constructs are any theory, hypothesis, or model that attempts to explain observed phenomena. The tasks that test-takers are required to perform must be logically samples of English language use in a defined context.

Q1:What is a proficiency test?

Q2:Is a proficiency test limited to one course or skill?

Q3:What language areas are usually tested in proficiency tests?

:Example :TOEFL

Q4:Are proficiency tests usually summative or formative?

Q5:Are proficiency tests norm-referenced or criterion-referenced?

Q6:How are the results of proficiency tests usually reported?

Q7:Why are proficiency tests often used for gate- keeping purposes?

Q8:Why do proficiency tests usually not provide diagnostic feedback?

Q9:Why should teachers not create their own proficiency tests?

Aptitude Tests

Finally, we need to consider the type of test that is given to a person prior to any exposure to the second language, A language aptitude test is designed to measure a person's capacity or general ability to learn a foreign language and to be successful in that undertaking. Aptitude tests are considered to be independent of a particular language . Language aptitude tests were ostensibly designed to apply to the classroom learning of any language.

There are two standardized aptitude tests were once in popular use:

A- The Modern Language Aptitude Test (MLAT) .

B- The Pimsleur Language Aptitude Battery (PLAB).

Both are English language tests and require students to perform such tasks as memorizing numbers and vocabulary, listening to foreign words, and detecting spelling clues and grammatical patterns. No research shows unequivocally that those kinds of tasks predict

communicative success in a language, especially untutored acquisition of the language. Because of this limitation, standardized aptitude tests are seldom used today.

Q1:What is an aptitude test?

Q2:What does an aptitude test aim to predict?

Q3:Are aptitude tests widely used today?

Q4:To what kind of learning were language aptitude tests designed to apply?

Q5:Name two standardized language aptitude tests used in the United States.

Q6:What type of tasks do aptitude tests usually include?

Q7:What do MLAT and PLAB measure?

Q8:What kind of success do MLAT and PLAB predict?

Q9:What limitation do standardized aptitude tests have?

Q10:Why are aptitude tests seldom used today?

Q11:What alternative approach is suggested instead of aptitude tests?

Q12:Why is any test claiming to predict language learning success considered flawed?

Q13:What does it mean to "pigeon-hole learners a priori"?

Q14:What should language learning presupposed?

Q/ What are the uses of assessment

1-Placement: to place or classify students on a suitable course.

2-Diagnosis: to specify student's particular strengths and weaknesses.

3-Selection: to select as fairly as possible the best candidates for a limited number of places or jobs or for higher education studies.

4-Evaluation : to evaluate the effectiveness of the syllabus as well as the method of teaching so as to make adjustments where needed.

5- Progress: to assess what students have learned of the specific syllabus and what progress they have made.

6-Prediction: to discover potential abilities and aptitudes and to predict probable future successes in school or outside.

7-Accreditation: to provide evidence that a student has undergone a course of training or study and is qualified to practice his profession.

Q/ Which one is wider (Assessment or test)?

Educational language assessment involves the use of a variety of procedures and tools to collect information about teaching and learning, and more importantly about students' language ability which is not only useful but necessary to prepare them for success. The most common kind of assessment procedure is testing .

Testing means presenting a student with a set of questions or tasks in order to obtain a measure of performance often represented by a score. The score is intended to give information about the person tested.

Q/ What are the differences between assessment and test

Test and assessment are used interchangeably but they do mean something different

Assessment	Test
1- Assessment is seen as a procedure instead of a product. 2- Assessment is used during and after the instruction has taken place. 3- Assessment is the systematic process of documenting and using empirical data to measure knowledge, skills, attitudes and beliefs.	1-Test is a "product" that measures a particular behavior or set of objectives. 2-Tests are done after the instruction has taken place, it is a way to complete the instruction and get the results. The results of the tests do not have to be interpreted, unlike assessment. 3-A test is used to examine someone's knowledge of something to determine what that person knows or has learned.

4- Assessment might show that students actually understands facts and figures, or a specific process. 5- Assessment can have different formats. A-They might require a student to answer questions . Or an assessment might be a teacher talking to a student about what they know. B-It could be a teacher's observation of a student working or talking about a subject. Assessment also be graded assignments, presentations, or class work that helps a teacher get an idea of what a student knows and doesn't know.	4-Testing might show the student's ability to memorize facts and figures, instead of a true understanding of those facts and figures. 5-Test usually follow a general format, where questions are asked and students answer them , they might be: -essay questions -multiple-choice questions -fill-in-the-blank questions -or true or false questions But the overall format is the same.
---	--