

University Batna 2  
Faculty of Letters and Foreign Languages  
Department of English Language and literature

Course: Testing and Evaluation  
Level: Master 1 LLA  
Teacher: Dr. N. Bentayeb

Semester 1 Course Content

<b>Chapter One: Testing</b> .....	<b>2</b>
<b>Lesson 1: What is Testing</b> .....	<b>2</b>
<b>Lesson 2: Testing, Assessment and Evaluation</b> .....	<b>4</b>
<b>Lesson 3: Reasons to Why Schools Use Tests</b> .....	<b>6</b>
<b>Lesson 4: The Purpose of Educational Testing</b> .....	<b>8</b>
<b>Lesson 5: Types of Tests</b> .....	<b>10</b>
<b>Lesson 6: Language Test Formats</b> .....	<b>12</b>
<b>Lesson 7: Testing Pros and Cons</b> .....	<b>17</b>
<b>Lesson 8: Characteristics of a Good Language Test</b> .....	<b>19</b>

## **Chapter One: Testing**

### **Lesson 1: What is Testing**

In general, testing is finding out how well something works. In terms of human beings, testing tells what level of knowledge or skill has been acquired.

A test or examination is an “assessment” intended to measure a test-taker's knowledge, skill, aptitude, physical fitness, or classification in many other topics.

A test may be administered orally, on paper, on a computer, or in a confined area that requires a test taker to physically perform a set of skills.

Tests vary in style, rigor and requirements. For example, in a closed book test, a test taker is often required to rely upon memory to respond to specific items whereas in an open book test, a test taker may use one or more supplementary tools such as a reference book or calculator when responding to an item.

A test may be administered formally or informally. An example of an informal test would be a reading test administered by a parent to a child. An example of a formal test would be a final examination administered by a teacher in a classroom or an I.Q (intelligence quotient) test administered by a psychologist in a clinic. Formal testing often results in a grade or a test score.

A test score may be interpreted with regards to a norm or criterion, or occasionally both. The norm may be established independently, or by statistical analysis of a large number of participants. A standardized test is any test that is administered and scored in a consistent manner to ensure legal defensibility. Standardized tests are often used in education, professional certification, psychology, the military, and many other fields.

#### **➤ Origin and Etymology**

From Middle English *test, teste*, borrowed from Old French *test, teste* (an earthen vessel, especially a pot in which metals were tried), from Latin *testum* (the lid of an earthen vessel, an earthen vessel, an earthen pot).

❖ **Noun:** test (plural tests)

a) A challenge, trial.

*“Numerous experimental tests and other observations have been offered in favor of animal mind reading, and although many scientists are skeptical, others assert that humans are not the only species capable of representing what others do and don’t perceive and know.” (Colin, 2012, p. 168)*

b) A cupel or cupelling hearth in which precious metals are melted for trial and refinement.

c) (academic) An examination, given often during the academic term.

d) A session in which a product or piece of equipment is examined under every day or extreme conditions to evaluate its durability, etc.

e) (Cricket, normally “Test”) A Test match.

f) (Marine biology) The external calciferous shell, or endoskeleton, of an echinoderm, e.g. sand dollars and sea urchins.

g) (Botany) Testa; seed coat.

h) (Obsolete) Judgement; distinction; discrimination.

### ❖ **Derived Terms**

- breath test
- crash test
- DNA test
- driving test
- final test
- fore test
- Pap test
- paternity test
- pregnancy test
- put to the test
- speaking test
- test card
- test case
- test data
- test drive
- tester
- test flight
- test run
- test subject
- test track
- test tube
- unit test

### ❖ **Verb:**

- a) To challenge.
- b) Climbing the mountain tested our stamina.
- c) To refine (gold, silver, etc.) in a test or cupel; to subject to cupellation.
- d) To put to the proof; to prove the truth, genuineness, or quality of by experiment, or by some principle or standard; to try. To test the soundness of a principle; to test the validity of an argument
- e) (Academic) To administer or assign an examination, often given during the academic term to (somebody).
- f) To place a product or piece of equipment under everyday and/or extreme conditions and examine it for its durability, etc.
- g) (Copulative) To be shown by test. He tested positive for cancer.
- h) (Chemistry) To examine or try, as by the use of some reagent. to test a solution by litmus paper

## **Lesson 2: Testing, Assessment and Evaluation**

Assessment and testing are often used interchangeably. What's the difference between assessment and testing in education? When developing instruction, it's important to know what the difference is between assessment, testing and evaluation.

When defined within an educational setting, assessment, evaluation, and testing are all used to measure how much of the assigned materials students are mastering, how well students are learning the materials, and how well students are meeting the stated goals and objectives. Although you may believe that assessments only provide instructors with information on which to base a score or grade, assessments also help you to assess your own learning.

Education professionals make distinctions between assessment, evaluation, and testing. However, all you really need to understand is that these are three different terms for referring to the process of figuring out how much you know about a given topic and that each term has a different meaning. To simplify things, we will provide the bellow first sight brief definitions:

- A **test** or **quiz** is used to examine someone's knowledge of something to determine what he or she knows or has learned. Testing measures the level of skill or knowledge that has been reached.
- **Evaluation** is the process of making judgments based on criteria and evidence.
- **Assessment** is the process of documenting knowledge, skills, attitudes and beliefs, usually in measurable terms. The goal of assessment is to make improvements, as opposed to simply being judged. In an educational context, assessment is the process of describing, collecting, recording, scoring, and interpreting information about learning.

Now let's examine the concepts in more depth:

### **What is an assessment?**

Assessment is the systematic process of documenting and using empirical data on the knowledge, skills, attitudes and beliefs. By taking the assessment, teachers try to improve student learning.

The verb assess comes from the French "assesser", but the origin is from the Medieval Latin "assessare" meaning "fix a tax upon". Another derivation of the Latin term is "assidere" or "adsidere" meaning "to sit beside" (a judge). Reference is made to the assistant of the judge whose job was to fix the amount of a fine or tax by estimating the value of a property.

Assessment is thus the process of collecting information about students from diverse sources so that educators can form an idea of what they know and can do with this knowledge. While evaluation is concerned with making judgments about instruction, a curriculum, or an educational system, assessment is concerned with the students' performance. In other words, one assesses an individual but evaluates a program, a curriculum, an educational system, etc.

***"Assessment occurs when judgments are made about a learner's performance, and entails gathering and organizing information about learners in order to make decisions and judgments about their learning."*** (Le Grange & Reddy, 1998, p.3)

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

Educators assess their students for a variety of purposes:

- To evaluate learners' educational needs.
- To diagnose students' academic readiness.
- To measure their progress in a course.
- To measure skill acquisition.

### **What is testing?**

What is testing in education? Almost everybody has experienced testing during his or her life. Grammar tests, driving license test etc. A test is used to examine someone's knowledge of something to determine what that person knows or has learned. It measures the level of skill or knowledge that has been reached. An evaluative device or procedure in which a sample of an examinee's behaviour in a specified domain is obtained and subsequently evaluated and scored using a standardized process (The Standards for Educational and Psychological Testing, 1999).

### **So, what is the difference?**

Test and assessment are used interchangeably, but they do mean something different. A test is a "product" that measures a particular behaviour or set of objectives. Meanwhile assessment is seen as a procedure instead of a product. Assessment is used during and after the instruction has taken place. After you've received the results of your assessment, you can interpret the results and in case needed alter the instruction. Tests are done after the instruction has taken place, it's a way to complete the instruction and get the results. The results of the tests don't have to be interpreted, unlike assessment.

### **What is evaluation?**

The verb "evaluate" means to form an idea of something or to give a judgment about something. The term comes from the French word "évaluer", meaning "to find the value of". The origin is from the Latin term "valere" meaning "be strong, be well; be of value, or be worth".

In the educational context, the verb 'to evaluate' often collocates with terms such as:

- the effectiveness of an educational system,
- a program,
- a course,
- instruction
- a curriculum

According to Weiss (1972), evaluation refers to the systematic gathering of information for the purpose of making decisions. It is not concerned with the assessment of the performance of an individual, but rather with forming an idea of the curriculum and making a judgment about it. This judgment is made based on some kind of criteria and evidence. The purpose is to make decisions about the worth of instruction, a course, or even the whole curriculum. Evaluation is thus larger and may include an analysis of all the aspects of the educational system.

### **Lesson 3: Reasons to Why Schools Use Tests**

Teachers teach content then test students. This cycle of teaching and testing is familiar to anyone who has been a student. Tests seek to see what students have learned. However, there can be other more complicated reasons as to why schools use tests.

At the school level, educators create tests to measure their students' understanding of specific content or the effective application of critical thinking skills. Such tests are used to evaluate student learning, skill level growth and academic achievements at the end of an instructional period, such as the end of a project, unit, course, semester, program or school year.

#### **➤ Testing assesses what students have learned**

The obvious point of classroom testing is to assess what students have learned after the completion of a lesson or unit. When the classroom tests are tied to well-written lesson objectives, a teacher can analyse the results to see where the majority of students did well or need more work. This information may help the teacher create small groups or to use differentiated instructional strategies.

Educators can also use tests as teaching tools, especially if a student did not understand the questions or directions. Teachers may also use tests when they are discussing student progress at team meetings, during student assistance programs or at parent-teacher conferences.

#### **➤ Testing identifies student strengths and weaknesses**

Another use of tests at the school level is to determine student strengths and weaknesses. One effective example of this is when teachers use pre-tests at the beginning of units to find out what students already know and figure out where to focus the lesson. There is an assortment of literacy tests that can help target a weakness in decoding or accuracy as well as learning style and multiple intelligences tests to help teachers learn how to meet the needs of their students through instructional techniques.

#### **➤ Testing identifies gaps in knowledge**

Taking a test permits students to assess what they know and what they do not know, so that they can concentrate study efforts on areas in which their knowledge is deficient. Students may take a test, realize which questions or items they got wrong, and then spend more time studying the items they missed. For example, Amlund, Kardash, and Kulhavy (1986) found that subjects corrected errors on a second test if they had an intervening study session after the first test. Other research shows that when students receive opportunities to restudy material after a test, they spend longer on restudying items that were missed than those that were correctly retrieved (Son & Kornell, 2008).

#### **➤ Testing causes students to learn more from the next study episode**

Another benefit of retrieval practice is that it can enhance learning during future study sessions. That is, when students take a test and then restudy material, they learn more from the presentation than they would if they restudied without taking a test. This outcome is called test-potentiated learning (Izawa, 1966). The benefits of test potentiation are distinctly different from the direct benefits of testing itself.

#### **➤ Testing produces better organization of knowledge**

Another indirect benefit of retrieval practice is that it can improve the conceptual organization of practiced materials, especially on tests that are relatively open-ended such as free recall in the lab or essay tests in the classroom. (Gates, 1917) postulated that one of the reasons retrieval practice leads to increased performance is that retrieval (or recitation, as he called it) causes students to organize information more than does reading. He suggested that as students actively recall material, they are more likely to notice important details and weave them into a cohesive structure.

#### **➤ Testing improves transfer of knowledge to new context**

Recent research shows that the mnemonic benefits of taking a test are not limited to the specific questions or facts that were tested; retrieval practice also improves transfer of knowledge to new contexts. Transfer may be defined as applying knowledge learned in one

situation to a new situation. Researchers often categorize transfer as being near or far; near transfer occurs if the new situation is similar to the learning situation, whereas far transfer occurs if the new situation is very different from the learning situation.

➤ **Testing can facilitate retrieval of material that was not tested**

One potential limiting factor of implementing testing in a classroom setting is choosing which material to test. It is unrealistic for an instructor to test students on everything. Fortunately, research on testing suggests that retrieval practice does not simply enhance retention of the individual items retrieved during the initial test: taking a test can also produce retrieval-induced facilitation; a phenomenon that shows testing also improves retention of non-tested but related material.

➤ **Testing improves metacognitive monitoring**

Another benefit of testing is improvement of metacognitive accuracy relative to restudying. Testing permits students to have better calibration of their knowledge. If students only study material repeatedly, they may think that their familiarity with the material means that they know it and can retrieve it whenever needed. However, such familiarity can be misleading. These points have direct implications for educational settings—the better students are at differentiating what they do know and what they do not know well, the better they will be at acquiring new and more difficult material and studying efficiently (Thomas & McDaniel, 2007). Therefore, instead of simply restudying, teachers can administer quizzes and students can self-test to determine what material they know well and what material they do not know well. Students' ability to accurately predict what they know and do not know is an important skill in education, but unfortunately students often make inaccurate predictions. When students reread material repeatedly, they are often overconfident in how well they know the material. Taking a test, however, can lead to students becoming less confident, a finding known as the under confidence-with-practice effect (Koriat, Scheffer, & Ma'ayan, 2002). Testing can help compensate for the tendency to be overly confident, which results in a more accurate assessment of learning.

➤ **Testing prevents interference from prior material when learning new material**

Another benefit of testing is that tests create a release from proactive interference. Proactive interference occurs when sets of materials are learned in succession; the previous material learned influences the retention of new materials in a negative manner. Thus, proactive interference refers to the poorer retention of material learned later, caused by prior learning (Underwood, 1957). Extended study sessions may therefore cause a build-up of proactive interference. However, research has shown that when tests are inserted between study episodes, they cause a release from proactive interference and enable new learning to be more successful.

➤ **Testing provides feedback to instructors**

Testing can provide teachers with valuable feedback about what students do and do not know. Tests and quizzes in the classroom are perhaps one of the most important ways in which teachers can formally assess the knowledge of their students, but of course homework can be used for this purpose, too. Testing is typically seen as an evaluation of what students have learned, and indeed this is true. Conscientious teachers will pay attention to how students perform on tests and use that knowledge to update their teaching in the future. If many students fail a particular topic on the test, it may be a sign to spend more time covering that material next time or use a different approach to teaching the materials and guide further instruction.

➤ **Frequent testing encourages students to study**

Having frequent quizzes, tests, or assignments motivates students to study. Every professor and every student knows that many students procrastinate and often do not study until the night before a test. Often university courses include only a midterm and a final exam, it is no surprise that the episodes of studying occur primarily just before tests. Mawhinney, Bostow, Laws, Blumenfeld, and Hopkins (1971) documented this point in controlled circumstances, with tests given daily, weekly, or every three weeks. Studying was most copious and evenly spaced with daily testing. With less frequent testing, study behaviour occurred only before the tests.

## **Lesson 4: The Purpose of Educational Testing**

Within university circles, testing refers to a broad variety of assessments. These range from informal questioning of students during lectures and tutorials to formal assessment such as quizzes, term papers, mid-term papers, practical work, continuous assessment tests, project work, terminal examinations and scientific research items used to assess skills, knowledge, attitudes and experiences. Formal, end-term examinations also fall within the scope of this broad definition of testing.

University examinations go back in history to the Middle Ages in Europe. With time, the form of tests employed, both at university and school level, has become more sophisticated. In the recent past, we have seen a gradual move away from exclusive reliance on terminal examinations to greater use of continuous assessment, special projects, research papers, portfolios of materials collected or work done during the course, along with terminal examinations.

Testing has long served a multitude of purposes in schools. According to Messick (1985, pp. 5-6) tests are used to:

- Assess student learning, both in terms of the achievement level attained and in terms of gains in achievement over the school year.
- Assess potential learning ability or readiness.
- Monitor progress.
- Diagnose reasons for failure.
- Judge degree of mastery.
- Place students in advanced or special programs.
- Provide a basis for school marks.
- Provide a basis for individualizing instruction.
- Identify under- and over-achievers.
- Guide students in their choice of specific school subjects and in their choice of curricula.
- Guide students in their decisions about postsecondary education and in their choice of specific colleges.
- Guide students in occupational and vocational choice.
- Inform institutions of higher learning about their applicants for admission.
- Inform prospective employers about job applicants.
- Inform students about their own abilities and achievements.
- Inform teachers about the abilities and achievements of their students.
- Certify minimum student performance for graduation.
- Assign students to sections in a course or to a grade by achievement or ability level.
- Meet state testing requirements.
- Select students into a school or program.
- Make school-to-school comparisons.
- Help in informing and counselling parents about children schooling progress.
- Evaluate programs and curricula, and to evaluate teacher effectiveness.



Compacting this list of testing purposes, we find that the major functions of testing in schools include:

- Assessing student abilities and achievements.
- Diagnosing learning difficulties.
- Grouping students for instruction.
- Placement in special or advanced programs.
- Informing students, teachers, and parents about student strengths and weaknesses.
- Counselling students and parents.
- Selection into schools or programs.
- Evaluation of program and teacher effectiveness.

In brief, tests are used for instructional improvement, certification of student achievement, placement, guidance, selection, and evaluation

These are traditional and familiar testing purposes but they are not exhaustive. For example, tests are used not only for student, teacher, and program evaluation but for system evaluation, monitoring, and reform as well. Tests are also used to clarify and operationalize educational objectives and to set realistic standards of excellence (Messick, 1985). Tests may also serve to assure equity of access to desirable programs and to needed student services (Heller, Holtzman, Messick, 1982).

## **Lesson 5: Types of Tests**

While assessment can take a wide variety of types in education, the following descriptions provide a comprehensive overview of a few major types of educational assessment.

**1. High-stakes Assessments:** are typically standardized tests used for the purposes of accountability—i.e., any attempt by state, or local government agencies to ensure that students are enrolled in effective schools and being taught by effective teachers. In general, “high stakes” means that important decisions about students, teachers, schools, or districts are based on the scores students obtain.

### **High Stakes vs. Low Stakes**

A “low-stakes test” would be used to measure academic achievement, identify learning problems, or inform instructional adjustments, among other purposes. What distinguishes a high-stakes test from a low-stakes test is not its form (how the test is designed) but its function (how the results are used). For example, if test results are used to determine an important outcome, such as whether a student receives a high school diploma, the test would be considered a high-stakes test regardless of whether it’s a multiple-choice exam, an oral exam, or an essay exam. Low-stakes tests generally carry no significant or public consequences - the results typically matter far more to an individual teacher or student than to anyone else - and scores are not used to burnish or tarnish the reputation of teachers or schools.

While high-stakes tests come in many forms and may be used for a wide variety of purposes, the following provide a brief overview of a few representative applications of high-stakes testing:

- **Students:** (learners) Test results may be used to determine whether students advance to the next grade level or whether they receive a diploma. For example, a growing number of states require students to pass a reading test to advance from third grade to fourth grade, while others require students to pass a test to graduate from high school.
- **Educators:** Test results may be used in the job-performance evaluations of teachers or to determine professional compensation. For example, in recent years more school reformers, elected officials, and policy makers have been calling for teacher pay (including bonuses), as well as hiring, firing, and tenure decisions, to be partly based on student test scores.
- **Schools:** Test results may be used to trigger penalties for schools, including negative public ratings, the replacement of staff members, or even closure. For example, some federal and state policies require that test results be used to impose a variety of consequences, such as firing or transferring some or all of a school’s administrators and faculty, or forcing a school to pay for additional services and transportation costs for students. In addition, standardized-test scores are also increasingly being used, along with other measures, in various state and independent efforts to assign A–F letter grades to schools.

**2. Pre-assessments:** are administered before students begin a lesson, unit, course, or academic program. Students are not necessarily expected to know most, or even any, of the material evaluated by pre-assessments—they are generally used to: (1) establish a baseline against which educators measure learning progress over the duration of a program, course, or instructional period, or (2) determine general academic readiness for a course, program, grade level, or new academic program that student may be transferring into.

**3. Formative Assessments:** are in-process evaluations of student learning that are typically administered multiple times during a unit, course, or academic program. The general purpose of formative assessment is to give educators in-process feedback about what students are learning or not learning so that instructional approaches, teaching materials, and academic support can be modified accordingly. Formative assessments are usually not scored or graded,

and they may take a variety of forms, from more formal quizzes and assignments to informal questioning techniques and in-class discussions with students.

**4. Summative Assessments:** are used to evaluate student learning at the conclusion of a specific instructional period- typically at the end of a unit, course, semester, program, or school year. Summative assessments are typically scored and graded tests, assignments, or projects that are used to determine whether students have learned what they were expected to learn during the defined instructional period.

Formative assessments are commonly said to be “**for**” learning because educators use the results to modify and improve teaching techniques during an instructional period, while summative assessments are said to be “**of**” learning because they evaluate academic achievement at the conclusion of an instructional period. Or as assessment expert Paul Black put it, “When the cook tastes the soup, that’s formative assessment. When the customer tastes the soup, that’s summative assessment.”

**5. Interim Assessments:** are used to evaluate where students are in their learning progress and determine whether they are on track to performing well on future assessments, such as standardized tests, end-of-course exams, and other forms of “summative” assessments. Interim assessments are usually administered periodically during a course or school year (for example, every six or eight weeks) and separately from the process of instructing students (i.e., unlike formative assessments, which are integrated into the instructional process).

**6. Placement Assessments:** are used to “place” students into a course, course level, or academic program. For example, an assessment may be used to determine whether a student is ready for Algebra I or a higher-level algebra course, such as an honors-level course. For this reason, placement assessments are administered before a course or program begins, and the basic intent is to match students with appropriate learning experiences that address their distinct learning needs.

**7. Screening Assessments:** are used to determine whether students may need specialized assistance or services, or whether they are ready to begin a course, grade level, or academic program. Screening assessments may take a wide variety of forms in educational settings, and they may be developmental, physical, cognitive, or academic. A preschool screening test, for example, may be used to determine whether a young child is physically, emotionally, socially, and intellectually ready to begin preschool, while other screening tests may be used to evaluate health, potential learning disabilities, and other student attributes.

**8. Diagnostic Assessments:** This testing is used to “diagnose” what skills a student has demonstrated proficiency on. Diagnostic testing often measures for student misconceptions or where students are in stages along a progression, such as by grade level, of concepts, or skills. For example, diagnostic reading assessments can measure what grade level students are fluent at reading, or based on their comprehension of the text.

Teachers use diagnostic testing information to guide what and how they teach. They’ll spend more time teaching skills students struggled with most on the diagnostic test.

Diagnostic assessments can be a very helpful tool for parents. The feedback kids receive on these tests lets the parents know the specific areas where they may need extra help at home.

## **Lesson 6: Language Test Formats**

Language test could be either written or oral.

**Written tests:** are tests that are administered on paper or on a computer (as an E-exam). A test taker who takes a written test could respond to specific items by writing or typing within a given space of the test or on a separate form or document.

**Oral tests:** The oral exam is a practice in many schools and disciplines in which an examiner poses questions to the student in spoken form. The student has to answer the question in such a way as to demonstrate sufficient knowledge of the subject to pass the exam.

A test developer's choice of which style or format to use when developing a written test is usually arbitrary given that there is no single invariant standard for testing. Be that as it may, certain test styles and formats have become more widely used than others. Below is a list of those formats of test items that are widely used by educators and test developers to construct paper or computer-based tests. As a result, these tests may consist of only one type of test item format (e.g., multiple choice test, essay test) or may have a combination of different test item formats (e.g., a test that has multiple choice and essay items).

Test items can be written in various formats, including **multiple choice, matching, true/false, short answer, and essay**. These formats vary in their strengths and weaknesses, and no one format is ideal in all circumstances.

The first three formats are known as **selected-response** formats, because the student sees the possible answers and has to choose (or select) the correct one.

The last two formats are known as **constructed-response** formats, because the student has to come up with the answers on his own.

### **1) Multiple Choice**

Multiple choice questions are composed of one question (stem) with multiple possible answers (choices), including the correct answer and several incorrect answers (distractors). Typically, students select the correct answer by circling the associated number or letter, or filling in the associated circle on the readable response sheet.

#### **Distractors are:**

- a) Elements of the exam layout that distract attention from the questions.
- b) Incorrect but plausible choices used in multiple choice questions.
- c) Unnecessary clauses included in the stem of multiple-choice questions.

Students can generally respond to this type of questions quite quickly. As a result, they are often used to test student's knowledge of a broad range of content. Creating these questions can be time consuming because it is often difficult to generate several plausible distractors. However, they can be marked very quickly.

➤ **Tips for writing good multiple-choice items:**

Avoid	Use
<p><b>In the stem:</b></p> <ul style="list-style-type: none"> <li>• Long / complex sentences</li> <li>• Trivial statements</li> <li>• Negatives and double-negatives</li> <li>• Ambiguity or indefinite terms, absolute statements, and broad generalization</li> <li>• Extraneous material</li> <li>• Item characteristics that provide a clue to the answer</li> </ul>	<p><b>In the stem:</b></p> <ul style="list-style-type: none"> <li>• Your own words – not statements straight out of the textbook</li> <li>• Single, clearly formulated problems</li> </ul>
<p><b>In the choices:</b></p> <ul style="list-style-type: none"> <li>• Statements too close to the correct answer</li> <li>• Completely implausible responses</li> <li>• ‘Overlapping responses</li> </ul>	<p><b>In the choices:</b></p> <ul style="list-style-type: none"> <li>• Plausible and homogeneous distractors</li> <li>• Statements based on common student misconceptions</li> <li>• True statements that do not answer the questions</li> <li>• Short options – and all same length</li> <li>• Correct options evenly distributed over A, B, C, etc.</li> <li>• At least 3 alternatives</li> </ul>

## 2) True/False

True/false questions are only composed of a statement. Students respond to the questions by indicating whether the statement is true or false.

Like multiple choice questions, true/false questions:

- Are most often used to assess familiarity with course content and to check for popular misconceptions.
- Allow students to respond quickly so exams can use a large number of them to test knowledge of a broad range of content.
- Are easy and quick to grade but time consuming to create.

True/false questions provide students with a 50% chance of guessing the right answer. For this reason, multiple choice questions are often used instead of true/false questions.

➤ **Tips for writing good true/false items:**

Avoid	Use
<ul style="list-style-type: none"> <li>• Negatives and double-negatives</li> <li>• Long / complex sentences</li> <li>• Trivial material</li> <li>• Broad generalizations</li> <li>• Ambiguous or indefinite terms</li> </ul>	<ul style="list-style-type: none"> <li>• Your own words</li> <li>• The same number of true and false statements (50 / 50) or slightly more false statements than true (60 / 40) – students are more likely to answer true</li> <li>• One central idea in each item</li> </ul>

**Suggestion:** You can increase the usefulness of true/false questions by asking students to correct false statements.

### 3) Matching

Students respond to matching questions by pairing each of a set of stems (e.g., definitions) with one of the choices provided on the exam. These questions are often used to assess recognition and recall and so are most often used in courses where acquisition of detailed knowledge is an important goal. They are generally quick and easy to create and mark, but students require more time to respond to these questions than a similar number of multiple choice or true/false items.

➤ **Tips for writing good matching items:**

Avoid	Use
<ul style="list-style-type: none"> <li>• Long stems and options</li> <li>• Heterogeneous content (e.g., dates mixed with people)</li> <li>• Implausible responses</li> </ul>	<ul style="list-style-type: none"> <li>• Short responses 10-15 items on only one page</li> <li>• Clear directions</li> <li>• Logically ordered choices (chronological, alphabetical, etc.)</li> </ul>

**Suggestion:** You can use some choices more than once in the same matching exercise. It reduces the effects of guessing.

### 4) Short Answers

Short answer questions are typically composed of a brief prompt that demands a written answer that varies in length from one or two words to a few sentences. They are most often used to test basic knowledge of key facts and terms.

An example this kind of short answer question follows:

**Q:** What do you call an exam format in which students must uniquely associate a set of prompts with a set of options?

**Answer:** Matching questions.

Short answer questions can also be used to test higher thinking skills, including analysis or evaluation. For example:

Will you include short answer questions on your next exam? Please justify your decision with two to three sentences explaining the factors that have influenced your decision.

➤ **Tips for writing good short answer items:**

Avoid	Use
<ul style="list-style-type: none"><li>• Trivia</li><li>• Long / complex sentences</li></ul>	<ul style="list-style-type: none"><li>• Your own words</li><li>• Specific problems</li><li>• Direct questions</li></ul>

**Suggestion:** When using short answer questions to test student knowledge of definitions consider having a mix of questions, some that supply the term and require the students to provide the definition, and other questions that supply the definition and require that students provide the term.

**5) Essays**

Essay questions provide a complex prompt that requires written responses, which can vary in length from a couple of paragraphs to many pages. Like short answer questions, they provide students with an opportunity to explain their understanding and demonstrate creativity, but make it hard for students to arrive at an acceptable answer by bluffing. They can be constructed reasonably quickly and easily but marking these questions can be time-consuming and grader agreement can be difficult.

➤ **Tips for writing good essay items:**

Avoid	Use
<ul style="list-style-type: none"><li>• Complex, ambiguous wording</li><li>• Questions that are too broad to allow time for an in-depth response</li></ul>	<ul style="list-style-type: none"><li>• Your own words</li><li>• Words like ‘compare’ or ‘contrast’ at the beginning of the question</li><li>• Clear and unambiguous wording</li><li>• A breakdown of marks to make expectations clear</li><li>• Time limits for thinking and writing</li></ul>

The difficulties with essay items are primarily administrative: for example, test takers require adequate time to be able to compose their answers. When these questions are answered, the answers themselves are usually poorly written because test takers may not have time to organize and proofread their answers. In turn, it takes more time to score or grade these items. When these items are being scored or graded, the grading process itself becomes subjective as non-test related information may influence the process. Thus, considerable effort is required to minimize the subjectivity of the grading process. Finally, as an assessment tool, essay questions may potentially be unreliable in assessing the entire content of a subject matter.

**6) Completion Type**

A fill-in-the-blank item provides a test taker with identifying characteristics and requires the test taker to recall the correct term. There are two types of fill-in-the-blank tests. The easier version provides a word bank of possible words that will fill in the blanks. For some exams all words in the word bank are used exactly once. If a teacher wanted to create a test of medium difficulty, they would provide a test with a word bank, but some words may be used more than once and others not at all. The hardest variety of such a test is a fill-in-the-blank test in which no word bank is provided at all. This generally requires a higher level of understanding and memory than a multiple-choice test. Because of this, fill-in-the-blank tests (with no word bank) are often feared by students.

## **7) Quizzes**

A quiz is a brief assessment which may cover a small amount of material that was given in a class. Some of them cover two to three lectures that were given in a period of times as a reading section or a given exercise in were the most important part of the class was summarize. However, a simple quiz usually does not count very much, and instructors usually provide this type of test as a formative assessment to help determine whether the student is learning the material. In addition, doing this at the time the instructor collected all can make a significant part of the final course grade.

## **8) Oral Exams**

Oral examinations allow students to respond directly to the instructor's questions and/or to present prepared statements. These exams are especially popular in language courses that demand 'speaking' but they can be used to assess understanding in almost any course by following the guidelines for the composition of short answer questions. Some of the principle advantages to oral exams are that they provide nearly immediate feedback and so allow the student to learn as they are tested. There are two main drawbacks to oral exams: the amount of time required and the problem of record-keeping. Oral exams typically take at least ten to fifteen minutes per student, even for a midterm exam. As a result, they are rarely used for large classes. Furthermore, unlike written exams, oral exams don't automatically generate a written record. To ensure that students have access to written feedback, it is recommended that instructors take notes during oral exams using a rubric and/or checklist and provide a photocopy of the notes to the students.



## Lesson 7: Testing Pros and Cons

Standardized testing has ignited a worldwide debate in the last few years (or decades), and many parents feel understandably concerned about their children being judged on the basis of tests that, in some cases, don't seem to reliably correlate with actual learning or with successful college and career outcomes.

Standardized testing seems to be one of those controversies that will never truly be figured out and solved. Students, parents, and educators could keep proving all of the negative things about these tests, but it does not look like there is any light at the end of the tunnel that will completely get rid of them.

### Pros

- **Accountability:** One of the strongest cases for standardized tests is that educators and schools are held accountable for their students' test scores. The scores are public records and can have serious consequences for teachers and schools that under perform.
- **Measurable analytics:** Without explicit test scores, comparison would not be possible. This would make it difficult for states and organizations to examine the performances of different schools.
- **Structure:** Standardized testing provides an established set of standards that all teachers and students work towards together. Benchmarks along the way provide a way to measure student performance over time.
- **Objectivity:** These tests are scored by computers or test-scorers with no vested interest in the students in order to remove bias.
- **Granular Data:** The data that standardized testing produces can be organized by specific criteria such as socioeconomic status, ethnicity, or special needs. This provides valuable insight into schools' strengths and shortcomings and can help develop programs targeting areas of weakness.
- **Establishes a universal standard for learning:** As the name implies, standardized testing creates some level of uniformity in the curriculum in every school from the same state. This ensures that students in all geographic locations are exposed to the same quality of material and are striving for the same learning goals.
- **Provides parents with more information:** Standardized testing offers a completely objective picture of areas where your student excels and areas where they need to improve. This helps teachers and parents to identify struggles a student might be facing on an annual basis so that they can catch up to their peers more quickly.

## Cons

- **Inflexibility:** Students who excel in classroom settings may not perform well on tests due to anxiety, unfamiliarity with the test format, family matters, health issues, language barriers, or a number of other common conditions. Standardized tests don't take personal issues into consideration.
- **Time-wasting:** Because of the huge emphasis placed on high test scores, teachers are forced to spend an inordinate amount of time teaching the specific topics guaranteed to come up on the test. This allows less time for creativity and personalized learning.
- **Doesn't measure progress:** A standardized test occurs once in the year. This does not provide any measurement of true progress. Many believe that a student's success should be based on their level of growth from the beginning to the end of the school year.
- **Stressful:** Teachers and students alike feel the stress of standardized tests. Teachers' jobs are at stake. Students' ability to go to college or even graduate are at stake. All based on a couple hours of intense pressure.
- **It only offers students one chance for evaluation:** While some students perfectly comfortable demonstrating their knowledge in a written test, other students panic from the excessive pressure. With annual standardized testing, there is only one opportunity to perform well. If a student cracks from the pressure associated with the test, he might not provide an accurate representation of his accumulated knowledge.
- **It can affect students' self-esteem:** If a hardworking student is struggling with the core principles in a certain subject despite his best efforts, underperforming on a standardized test can make him feel like a failure.
- **It restricts creativity for teachers:** Now that teachers are evaluated based on their test scores, they must alter their lesson plans to focus on the hyper-specific skills required for one particular test. This often asphyxiates dynamic teachers who want to plan engaging, hands-on lessons for their classes each day.
- **Politics:** With public and charter schools competing for funds, standardized testing becomes a heavy factor. Many people claim that politicians use standardized test scores to further their own political agenda.

## **Lesson 8: Characteristics of a Good Language Test**

What are the characteristics of a good language test? A good test should have a positive effect on learning, and teaching should result in improved learning habits. Such a test will aim at locating the specific and precise areas of difficulties experienced by the class or the individual student so that assistance in the form of additional practice and corrective exercises can be given. The test should enable the teacher to find out which parts of the language program cause difficulty for the class. In this way, the teacher can evaluate the effectiveness of the syllabus as well as the methods and materials he or she is using. A good test should also motivate by measuring student performance without in any way setting "traps" for them. A well-developed test should provide an opportunity for students to show their ability to perform certain language tasks. A test should be constructed with the goal of having students learn from their weaknesses. In this way a good test can be used as a valuable teaching tool.

Therefore, a good language test should be:

### **1- Valid:**

It means that it measures what it is supposed to measure. It tests what it ought to test. A good test which measures control of grammar should have no difficult lexical items.

### **2- Reliable:**

If it is taken again by (same students, same conditions), the score will be almost the same regarding that the time between the test and the retest is of reasonable length. If it is given twice to same students under the same circumstances, it will produce almost the same results. In this case it is said that the test provides consistency in measuring the items being evaluated.

### **3- Practical:**

It is easy to be conducted, easy to score without wasting too much time or effort.

### **4- Comprehensive:**

It covers all the items that have been taught or studied. It includes items from different areas of the material assigned for the test so as to check accurately the amount of students' knowledge.

### **5- Relevant:**

It measures reasonably well the achievement of the desired objectives.

### **6- Balanced:**

It tests linguistic as well as communicative competence and it reflects the real command of the language. It tests also appropriateness and accuracy.

### **7- Appropriate in difficulty:**

It is neither too hard nor too easy. Questions should be progressive in difficulty to reduce stress and tension.

### **8- Clear:**

Questions and instructions should be clear. Pupils should know what to do exactly.

### **9- Authentic:**

The language of the test should reflect everyday discourse.

### **10- Appropriate for time:**

A good test should be appropriate in length for the allotted time.

### **11- Objective:**

If it is marked by different teachers, the score will be the same. Marking process should not be affected by the teacher's personality. Questions and answers are so clear and definite that the marker would give the students the score he/she deserves.

### **12- Economical:**

It makes the best use of the teacher's limited time for preparing and grading and it makes the best use of the pupil's assigned time for answering all items. So, we can say that oral exams in classes of +30 students are not economical as it requires too much time and effort to be conducted.