Assess Learner Knowledge

INTRODUCTION

How will you know whether your learners have acquired the technical knowledge implied or stated in the learner performance objectives for the program? How can you be sure that they have learned the required theory, facts, data, or other information? The answer is that you can assess their knowledge level using some type of cognitive test items: true-false, multiple-choice, matching, completion, essay, or oral items.

Of course, measurement of learner knowledge is only one part of the picture. In fact, there are three kinds of performance-knowledge, skills, and attitudes-that you will want to measure, and these are not truly discrete. In performing a skill, for example, a learner is not only demonstrating psychomotor skill, but also certain attitudes (e.g., concern for safety and cooperation) and the possession of certain knowledge (e.g., the proper procedure to follow in making change for a customer).

However, there are many times when the measurement of knowledge alone is important. You may want to determine whether learners know the steps in a procedure before allowing them to carry out that procedure for the first time. You may need to determine if learners have understood a particular reading assignment or class lecture. You may want to test their ability to solve practice problems. And so on.

By assessing learner knowledge regularly, you can keep them informed concerning their progress in learning the technical knowledge required for the occupation. You can determine their readiness to go on to subsequent learning activities. Assessment can also provide you with valuable information about the effectiveness of your own teaching-whether your learners are learning, or whether you need to change your teaching strategies. And assessment also provides a basis for assigning grades in traditional vocational-technical programs.

Assessing learner cognitive performance, under one name or another, is as old as education. Experts agree that it is one of the instructor’s most important responsibilities. For information on why, when, and how to assess your vocational learners’ cognitive performance, read the following information sheet.

COGNITIVE PERFORMANCE ASSESSMENT

Instruction should be based on identified learner performance objectives, which state the intended outcomes of the educational process in terms of the specific knowledge, skills,
and attitudes learners need for entry into the world of work.' Learning in all three
domains—knowledge (cognitive), skills (psychomotor), and attitudes (affective)—is
important, but in this module we are concerned with the cognitive domain.

Benjamin Blooms divides the cognitive domain into six levels. Cognitive objectives at the
lowest level (i.e., knowledge) require learners to recognize or recall correct facts, data,
or information. Such objectives usually call for learners to list, define, identify items, or in
some other way recognize or recall a particular piece or body of information.

Learner performance objectives at the higher levels in the cognitive domain require
learners to do more than simply recognize or recall factual information correctly. They
also require learners to use that information in some way. Objectives at the second level
of the cognitive domain (i.e., comprehension), for example, may call for learners to
summarize, interpret, translate, or paraphrase facts, data, or information. Objectives at
still higher levels may require application, analysis, synthesis, or evaluation of factual
information.

Learner cognitive performance assessment is the process of determining whether
learners have achieved these cognitive performance objectives. It is the process of
testing learners' knowledge and their ability to use that knowledge in a manner
appropriate to your occupational area.

**Purposes of Assessment**

Information about learners' cognitive performance—what knowledge they have acquired
and how they can use that knowledge—can help in providing high-quality instruction for
the learners enrolled in your program. You and other people can put this information to
several important uses.

**Learner readiness.** You can use this information to determine whether your learners
are ready for a particular learning activity. Often, learners must know one thing before
they proceed to another activity. Let's say, for example, that a home economics teacher
is demonstrating to learners how to separate egg yolks from whites. The teacher has
finished showing them the procedure and is now ready to have them try their skill on
some real eggs.

Before the teacher turns learners loose on those eggs, however, he or she might want to
make sure that they're ready for the task. Eggs do cost money and they can make a
nasty mess if mishandled. The teacher would be likely to quickly test learners' knowledge of the procedure—their readiness for the hands-on task. He/she would want to
ensure that learners know (1) the steps in the procedure (e.g., first you break the egg);
and (2) the critical factors in any steps (e.g., you must break the egg carefully, so that
the two halves of the shell remain intact).

For that matter, you will undoubtedly test learners' knowledge of safety rules and
practices before learners actually begin to use the tools and equipment in the laboratory.
And, you will often test learners' knowledge of theory before they proceed to practical
exercises using that theory. Perhaps carpentry learners are ready to begin construction
projects including the purchase of lumber. The carpentry teacher would probably want to
test learners' knowledge of formulas for converting running feet to board feet before they make their purchases.

**Instructional improvement.** You can also use the information you get through assessing learner cognitive performance to improve your instruction. Let's take the egg-separating example again. The teacher might test learners' knowledge of the procedure and discover that they don't know what he/she thought they did. This could indicate that the instructional activities need improvement. Perhaps the demonstration was unclear or confusing. Perhaps some learners were just sitting too far away to see what the teacher was doing. Finding out that the learners haven't learned what you were teaching can alert you to possible weaknesses in your instruction.

**Information about learner progress.** Other people can make good use of this information about learners' learning, too. Learners, for instance, need to know how they are progressing in your program. They will want to know if they are successfully acquiring the technical knowledge they will need on the job.

Providing learners with feedback concerning their cognitive performance can help them identify their own weaknesses so that they can work more successfully to acquire needed technical knowledge. Positive feedback on the progress they are making can help to motivate them to progress further.

Administrators will also be interested in the information you gain by assessing learner cognitive performance. One of your responsibilities, as a vocational-technical instructor will no doubt be to report learners' progress or achievement. Almost all vocational teachers must submit learner grades or other progress reports at specified intervals. Part of learners' overall progress, of course, is their progress in acquiring needed technical knowledge.

The information about learner progress that you report to administrators may also be shared with other people in the school or community. Counselors might use this information in making placement decisions with learners. Parents of secondary learners are usually concerned with how their children are doing in the vocational-technical program.

Finally, employers frequently want to know about your learners' competence. They commonly want to know learners' course grades. They might even want to talk to you about the technical content of your program or about the knowledge, attitudes, and skills of a particular learner. Of course, you can't tell potential employers what your learners know, and what they can do with what they know, until you have found out yourself by assessing their cognitive performance.

**Assessment Techniques**

You might assess your learners' cognitive performance at almost any time. It would not be exaggerating to say that you will assess cognitive performance, in one way or another, for one purpose or another, every day in your classes. Ensuring that learners have mastered the technical content of your program should be an ongoing part of your instruction.
There are a number of different ways to assess learner cognitive performance. In some situations, you might use very informal means to do the job. You might, for example, give a short oral quiz, requiring learners to provide brief written answers that you can quickly check, during class, before they move into the laboratory. A final course examination, on the other hand, would almost certainly involve a full blown, formal written test.

Regardless of how formal or informal the testing situation is, however, the test items that you will use to measure learner knowledge will be of two types: objective and subjective.

Some kinds of cognitive test items are called objective by testing experts because scoring them is almost entirely an objective process. The following types of items are considered to be objective:

- Multiple-choice
- Matching
- Completion
- True-false

Scoring these items is primarily objective because the scorer (i.e., you) does not need to use judgment to determine whether a learner's answer is right or wrong. Learners' answers can be compared to a simple scoring or answer key. If the key says that the answer to an item should be "true," a learner's answer either does or does not agree. Neither opinion nor judgment are needed for scoring.

Other kinds of test items are called subjective because they do require the use of judgment and interpretation in scoring answers. The following types of items are considered to be subjective:

- Essay
- Oral

As you are no doubt aware, scoring essay items is not as simple a matter as scoring objective items. Is a given answer the right answer? Is it partially right? If an essay question is worth 20 points, how many points is a partially right answer worth? Scorers must use their own judgment and opinion to resolve these questions in scoring subjective items.

There is another basic difference in test items. Some items require learners to supply the answer from their own memory. The subjective types of items are both supply tests, as you certainly know if you've ever taken an essay test. Of the objective types of test items, only one is of the supply type-completion items. Completion items require learners to fill in the blank or give a short answer (word, numeral, symbol, or phrase) to a question, again from their own memory.

The other types of objective test items (multiple-choice, matching, and true-false), however, are selection types. In all of these, learners are given two or more possible answers from which to select the correct answer or answers. Selection types test
learners' ability to recognize correct information, while supply types test learners' ability to recall correct information.

Finally, a word about terminology. You may have noticed in the previous explanation that we talked about test items rather than test questions and about multiple-choice items rather than multiple-choice tests. The reasons that the word item is the preferred word in both cases are as follows:

- Items on a test may or may not be stated in interrogative-question-form. True-false items, for example, are not questions but declarative statements.

- A given test may include many different types of items. While you might devise a test that included, for example, only essay items (i.e., an essay test), you will also devise tests that include one section of essay items, one of true false items, one of multiple-choice items, or some other combination.

Consequently, the word item is a more precise descriptor in both situations.

The Good Test: In Theory

What makes tests good or bad? What qualities should your tests have? Objective vs. subjective items, supply vs. selection items—how do you choose which to use? There is a wealth of literature in which testing experts agree that a good test should have the following characteristics.

The test must be valid. Validity is the extent to which a test measures what it is supposed to measure. To the extent that a test measures what it is supposed to, it is valid. This notion seems so simple and straightforward that you might wonder why it is even mentioned. How could you possibly go wrong?

You might go wrong if your test measures more than you intended. When you assess learner cognitive performance, you want your tests to measure learners' knowledge of specific facts, data, or information. Your test becomes less valid if it also measures other knowledge as well.

Assume, for example, that it is very important that learners have read an assigned chapter before moving on to the next learning activity. You thus want to administer a test to ensure that they have read and understood the assigned material. One of your test items is the following multiple-choice item:

What country is currently pioneering work in the area of friction welding?

a. India  
b. Greenland  
c. Nicaragua  
d. Russia

Some learners would be able to answer this item correctly using general knowledge, whether they had read the assigned material or not. Some learners would be likely to know that, of the four countries listed, the Soviet Union has by far the most
technologically advanced industrial economy. The Soviet Union would, therefore, be the most likely of the four countries to be engaged in pioneering work in a technological area. A guess based on this reasoning would be correct.

This item, then, measures learners' general knowledge as well as their knowledge of the reading assignment. Could the item be repaired to measure what it is supposed to? If all the possible answers were equally plausible (i.e., all countries listed had technologically advanced industrial economies), general knowledge would be useless. Learners would need to know the material covered in the assigned reading in order to answer the item. The item would then measure what it is supposed to, thereby increasing its validity.

Another way in which validity can be affected is if items on a test are answered by other items. For example, consider the following two completion items:

Harmful ___________ live and grow on inanimate objects. When harmful germs get inside a human body, ___________ can begin.

The first correct answer—germs—is provided in the second item. Learners who are discerning enough to recognize this can do well on the test without having truly learned the material. Thus, test validity is reduced.

The validity of a test can also be reduced if learners must use skills other than those you intend to test in order to answer the items. For example, any written test requires learners to use reading skills. Essay tests require the use of writing skills. Oral tests require the use of oral communication skills.

The need for learners to use these other skills can be a serious problem, causing reduced validity, if some learners cannot read or understand the items to begin with. If this happens, your test is measuring two things whether you intend it to or not—learners' technical knowledge and their communication skills.

Learners will need some level of communication skills to succeed in the world of work. We are not suggesting, by any means, that such skills are not important. But if you wish to measure technical knowledge, then you must ensure that your test allows learners to show how much technical knowledge they actually possess. If a learner's communication skills are weak, that problem should be identified and remediated, but it should not be allowed to cloud the measurement of technical skills.

The problem can be avoided if your items are carefully developed with learners' communication skill levels in mind. In other words, written items should be at learners' reading level. All items should be clearly and simply stated so that learners' communication skills are sufficient to allow them to demonstrate their technical knowledge. In this way, the effects of communication skills on validity will not be significant.

The length of a test can also affect its validity. Let's say, for example, that you are preparing a final examination for your learners covering the entire semester's technical content. You intend to test learners' acquisition of a considerable body of technical knowledge and their ability to use that knowledge. Your final examination, then, must be long enough to cover all the knowledge and use of knowledge that you intend to test.
If your final examination has only five multiple-choice items, you are probably only testing learners’ knowledge of five small bits of information and not of the whole of the technical content covered during the semester. The five-item final examination thus would have very low validity. If the test is to measure knowledge of the many bits that comprise the whole technical content, it has to be long enough to include a representative sampling of all those bits.

On the other hand, a five-item test could have very high validity in a different situation. Perhaps you want to determine if learners are ready to go on to hands-on activities after information activities. A five-item test might be quite sufficient if there are only five bits of information to check on. Since, in this case, you are sampling a small body of technical content, a short test should be sufficient.

The test should be reliable. Reliability is the consistency with which a test measures achievement. Experts in the theory of testing tell us that a valid test is always reliable—that is, it consistently measures what it is supposed to measure. Some theorists, in fact, treat reliability as a part of validity. In any case, it is important to note that a test can be reliable (i.e., can get a consistent measurement) even if it is not valid (i.e., does not measure what it is intended to measure).

This point can be illustrated with an extreme example. Let’s say that you decide to measure learners’ knowledge of some information. For some unexplained reason—perhaps you’ve had a very hectic day—you write your test in German, even though none of your learners understand a word of German. Now, obviously, this test would not be valid for any learner who didn’t speak German. Thus, it wouldn’t be valid for your learners. It would not measure what it is supposed to; instead, it would measure learners’ knowledge of the German language.

In spite of this, however, the test could be reliable. Let’s say that you actually administer this German language test to your learners. There is a good chance that all of them would score zero. Furthermore, they would get this same score consistently. You could give them the same test again the next day and they would still score zero (unless they all spent the previous evening closeted with a German tutor). Hence, although this German-language test would have no validity for English speakers, it would yield reliable results—consistent scores.

One factor that has a great effect on test reliability is the subjectivity of the scoring. If a test has high reliability, you should be able to administer and readminister it to a group of learners and get roughly the same set of scores (making allowance for such things as increased learning in the intervening time). Or, two different scorers should be able to sit down with a set of learner tests, rate them independently, and arrive at very similar scores.

Subjective scoring, however, can lower the reliability of tests because it can allow inconsistent measurement. Two different teachers could each independently score learner essay items and end up with completely different scores for the same answers. For that matter, one scorer could score a test one way one time yet score it completely differently another time.
Psychological studies have confirmed time and time again that inconsistent scoring can be a serious problem with subjective test items. Scorers must constantly make decisions—on the basis of their personal judgment, opinion, and preference—about the worth of learners' answers.

Thus, it is important to objectify subjective test scoring insofar as possible. For example, before administering an essay test, you need to determine whether points will be deducted for incomplete sentences or for spelling errors. (Are these skills part of what you wish to measure? How many points will you deduct for each error or as a maximum?) You also need to prepare a scoring key, which lists the key points you expect to be made in each answer, and you need to determine how much value to assign to each key point covered. A thorough scoring key can minimize reliability problems with subjective tests.

Objective tests, on the other hand, do minimize or even eliminate subjectivity in scoring (hence, their name). If a true-false item is well written, for example, there can be only one answer to it. Further, no interpretation or judgment is required on the part of the scorer to determine if a learner's answer is correct. The item is either true or false, and the learner either did or didn't mark the correct response.

Different scorers, consequently, should have no difficulty in arriving at the same score for an individual learner's test. Or, you could score the same set of papers twice and get exactly the same set of results. For that matter, with a scoring key, anyone should be able to score an objective test consistently, whether he/she knew one bit of the content tested or not.

As important as reliability is, however, it must be balanced with other concerns. Objective test items, in general, are good at measuring lower-level cognition (e.g., recall of facts). Higher-level cognition is more effectively measured through subjective test items. Thus, you must not, in the interests of reliability (and ease of scoring), use only objective measures. Rather, as mentioned before, you must ensure that your procedures for scoring subjective test items are as objective as possible.

The test should be usable. No matter how valid or reliable a test might be, it will be of little use to you if it is unreasonably difficult or time-consuming to prepare, administer, or score. A final examination with 1,000 objective items could be quite valid and reliable, for instance. Yet such a test would take so long to prepare, administer, and score, that it simply would not be practical to use.

Likewise, essay or oral items can in theory be used to test learner recall of facts, data, or information—the lowest level of the cognitive domain. However, essay and oral items can be quite time-consuming. It takes learners a longer time to write essay answers and you a longer time to score them. You can't give an oral item to the whole class at once (would they all answer together, in unison?), so considerable time would be needed for individual administration and scoring.

Consequently, essay and oral items would not be the most practical, usable means of assessing learner cognitive performance at the recall level. Much more practical and usable in this case would be objective items.
Other factors may affect the usability of tests as well. Ultimately, you will need to use your own judgment to determine any other qualities that would make a test practical and usable in your own situation.

The Good Test: In Practice

How can you develop tests that are valid, reliable, and usable in your own program? The following guidelines can help you ensure that your tests measure what they are supposed to measure, and measure it consistently, while remaining practical and usable.

Base your test on learner performance objectives. Just as your learner performance objectives identify the knowledge, skills, and attitudes to be taught, they also establish a basis for learner evaluation. Logically, a given test should cover each of the learner performance objectives that learners should have achieved at a particular point. As you identify cognitive objectives to be covered in your test, you can also identify the specific criteria, described in the objectives, against which learners' achievement should be measured.

You can then develop test items that also require learners to demonstrate the required knowledge at the specified level (e.g., recall, synthesis). This will be of great importance in ensuring that your test has high validity—that it actually measures what it is supposed to.

The test should differentiate. If your test measures what it is supposed to measure, it should differentiate between learners who know the material being tested and those who don't. If it does not differentiate, its validity will be low.

To tell if a test differentiates in this way, you will need to look at other indicators of learner knowledge or achievement. These could be general in nature—the grades learners usually get in your program, for example. If learners who usually get good grades do well on a test, while learners who usually get poor grades do badly on the test, it is probably reasonable to conclude that the test differentiates. Other general indicators of knowledge or achievement that you can use for comparison might be scores on standardized tests, general intelligence tests, or other tests of your own that you know to have high validity and reliability.

You should keep in mind, however, that many of these general indicators should be taken with a grain of salt. Standardized achievement tests and intelligence tests are often criticized by testing experts because of their cultural and language bias. Most appropriate for your purposes would be other specific indicators of knowledge of the particular content being tested.

For example, you might be able to compare learners' test scores with their scores on daily quizzes covering the same content. Your own observation of learner performance in the classroom or laboratory can often tell you which learners know the material and which don't. For example, a learner who does well on a test, and then also does well on hands-on practice activities that require using that specific knowledge, can reasonably be assumed to possess that knowledge.
Minimize the effects of communication skills. We have already discussed that every type of test item requires learners to use some communication skills to understand and answer the items. A test becomes less valid if it requires learners to use communication skills they don't have. When this happens, learners cannot demonstrate their knowledge because they cannot communicate it to you. The test is not measuring what it is supposed to.

Thus, you should choose your test items carefully. For example, the essay item probably requires the highest level of communication skills. If a cognitive performance objective requires that learners actually organize and present information themselves, then essay items should, of course, be used. But if simple recall is required, you can use one of the objective types of items that does not require such high-level communication skills.

You can also minimize the effect of communication skills by writing clear, simple test items. Everyone is familiar with the trick question—the kind that you have to read and interpret in one particular way, and no other, to get the right answer. Questions like this, however, simply favor learners who have greater skills in logic and in using the English language.

The more of these two skills the learner has, the greater is the chance that he/she will successfully decode the trick item, divine your intention, and answer correctly. If your test is to measure what it is supposed to, your questions should be simple enough to be read in the same way by all learners and answered correctly by all those who know the material.

This does not mean that you cannot test for fine distinctions in learners' knowledge. On the contrary, you can and sometimes should do so. If your occupational area is one in which fine distinctions are important, you will certainly want to test learners' knowledge of them. Furthermore, testing for these fine points can help your tests to differentiate between those who know the material and those who don't. Remember, nonetheless, that you can test for fine points in a clear, simple, direct manner.

The need to use communication skills in answering test items is a matter of particular importance for learners with exceptional needs. A learner with limited English proficiency, for example, might have considerable difficulty in understanding your test items or in writing answers to them. It is quite possible that such a learner may know the information being tested perfectly well but be unable to communicate his/her knowledge to you.

Similarly, a visually impaired learner might have difficulty reading a written test. A hearing-impaired learner could have trouble with an oral test. And a physically impaired learner might have problems in fitting written responses in the spaces provided. For these learners, special modifications may be needed in the testing process.

Provide clear, full, and simple directions. Difficult, complicated, incomplete directions can also reduce the validity and reliability of your test. If the directions are so difficult that some learners cannot read them, validity will be reduced because of the effects of communication skills.
If the directions are unclear or incomplete, reliability will be reduced because the measurement won't be consistent. Mary Lou will answer an item this way because she thinks she is supposed to, while Johnny will answer it a completely different way, again because he thinks he's supposed to. Should this occur, learners' scores will not truly indicate whether they knew the material or they didn't.

One good way to ensure that your test directions are clear, full, and simple is to test them out. Give the directions and a sample item to a learner who is not going to be taking the test. Ask the learner to read the directions and answer the item accordingly. You may want to test your directions with several different learners, just to be on the safe side. Once you have developed and tested your directions for a particular kind of test item, you can reuse those directions for other items of the same kind.

In addition, it is often a good idea to include a sample item, with the correct answer appropriately marked, along with your test directions. An example provides learners with another opportunity to be sure they understand your directions and answer appropriately. This is especially necessary if learners are not used to the kind of test item, the way you want answers marked on their papers, and so on.

Your test directions should first explain administrative details pertaining to the whole test. You should tell learners how much time is allowed for the test; how much time they should spend on various parts; and the point value of the test, the parts, and the individual items in the parts.

Learners will also need to know whether they are to mark their answers on the test paper or on a separate answer sheet and exactly how to mark their answers. You should tell them whether they are to write the appropriate word in the blank in each item, circle the number of the correct response, place an $X$ over the $T$ in the column if the statement is true, or whatever. You should give separate directions for marking answers, with examples if appropriate, for each different kind of item that you use on the test.

Learners should also be told what to do during the test if they should have additional questions. Their questions should be answered, but in such a way that it does not disturb the rest of the class. For example, you could ask learners to raise their hands should they have questions, so that you can go to where they are sitting and answer such questions on an individual, and quiet, basis.

Do not use too many different kinds of items. Authorities generally agree that you should use no more than three different kinds of items (e.g., true-false, essay, matching) on a single test. In addition, all items of the same kind should be grouped together.

Learners need to develop a particular mindset in order to answer a particular kind of item. Achieving that mindset may be difficult and time-consuming for some learners. You shouldn't, therefore, require learners to change their mindset more often than is necessary-unless you also want to measure their ability to change modes of reasoning and answering rapidly on command.

If you find that you have prepared a test with too many kinds of items, you have two possible solutions, both of which are simple and easy. First, you might rewrite some of the items to convert them from one type to another. You could include these revised
items with others of the same type, ending up with fewer kinds of items. Second, you could split your test into two tests and give learners the two tests on two different days.

And note, please, that when you use different types of items on a single test, you should number all items consecutively (1, 2, 3, and so on) from beginning to end. Do not start numbering items in each new section with the number 1. When you are discussing the items with learners, either during or after the test, it is much simpler if there is only one item per number. Otherwise, there can be confusion—"Which number 1 do you mean, the essay item or the true-false item?"

The test should be just the right length. Determining the right length for a test involves a compromise between validity, on the one hand, and reliability and usability, on the other. We discussed previously how the validity of a test depends on how comprehensively the content is sampled. The larger the body of knowledge to be tested, the lengthier the test should be if it is to be valid.

In some situations, then, you might naturally be tempted to prepare very long tests. A final examination, for example, might test for knowledge of large quantities of information. You could logically conclude that the best approach is to use a large quantity of items. The more comprehensive the sample, the more valid the test-right?

That is true. However, if the test is too long, reliability can suffer. The longer the test, the more likely it is that learners will become bored or fatigued. Fatigue and boredom are two learner variables that affect reliability, because they affect the consistency of the learners' performance in test-taking. If learners are tired or bored, they will be less likely to remember what they know. They may confuse their memories of one bit of knowledge with another.

You will need to use your own judgment and your knowledge of your learners, their needs, and their abilities to determine the right length for a given test. You should know your learners well enough to know approximately at what point fatigue and boredom will begin to adversely affect their test-taking performance.

Remember, too, that the best compromise on length often depends on the situation. The test must be long enough to be valid for your purposes, yet not so long that reliability is reduced.

Watch for the guessing factor. There is always the possibility that a learner could take a blind guess at any test item and get the right answer out of thin air. If a multiple-choice item has four choices, learners have a 25 percent chance of getting the correct answer without even reading the item.

Learners have a 50-50 chance on true-false items, which only offer two choices. Even in a matching item listing five inventions and the names of five inventors, the chances are only one in five—and sometimes the last item is free, by process of elimination.

One kind of guessing that you can and should prevent, however, is reasoned guessing. When you make a reasoned guess—an educated guess—you use whatever clues you can find in the item, together with your general knowledge, to figure out what the answer should be.
Multiple-choice items, for instance, often allow learners to guess the right answer by process of elimination—"Well, I know for sure that it's not A, and I've never heard of B before, and C just doesn't sound right, so it must be D." To minimize the success of educated guesses, you will need to ensure that each of your items is carefully devised and free of clues to the correct response.

Now, in some situations, you may want to ensure that learners cannot guess the correct answer because it is essential that they know the information being tested. Let's say that you're testing learners' knowledge of safety rules and practices at the beginning of the year or term. You want to be absolutely sure that learners know their safety rules and practices before they try using any tools, equipment, or machinery. In this situation, if they don't know, they might injure themselves.

The best approach when you need to be sure that they know is to avoid selection items. This means that you will have to decide whether to use essay items, oral items, or completion items. Completion items would be the most likely candidate. Your test can be valid, reliable, and usable—and learners cannot guess.

Make good copies for everyone. If you are administering a written test, each learner should have his or her own copy. So, be sure that you have enough well-produced copies of the test—neat, clean, and free from grammatical or typographical errors.

A sloppy production job (e.g., misspelled words, incomplete sentences, poor copies) invites learners to approach the test haphazardly. It also creates one more opportunity for learners to misunderstand your test directions or items.

Create a favorable testing environment. The physical environment in your classroom or laboratory should allow learners to concentrate on taking the test. The room should be reasonably quiet and free from distractions. The temperature should be comfortable, and lighting should be adequate. Ventilation should be sufficient to keep the air in the room fresh.

The testing environment includes psychological factors as well. Some learners become very anxious and apprehensive when taking a test, regardless of whether they know the material being tested or not. Or, learners may be apprehensive because yours is a new kind of test for them—perhaps you're giving them their first essay items. You will often be able to minimize such apprehension by making the testing environment nonterrorizing.

One way to do this is to make sure that learners understand the purpose of the test. Is the test to be graded, and are grades to be recorded? Or is the test an ungraded review of material? You should also tell learners in advance exactly what material is to be covered in the test. Otherwise, some learners will undoubtedly waste time trying to decide what to study—becoming, in the process, more and more apprehensive about the testing situation.

Another way to make the environment nonterrorizing is to give learners a practice test. This may be especially important if you are using a type of test item (e.g., essay) that your learners are not used to. Practice tests give learners the opportunity to gain experience with the type of test item and see that they can be successful at it.
Many testing experts also recommend putting easier questions at the beginning of your test. Their logic is that, since all learners should be able to answer these easier questions, all learners should be able to experience some initial success in taking the test, thereby decreasing their anxiety.

You will probably never be able to eliminate all unfavorable psychological factors in test-taking. You cannot control what might happen to your learners outside your own classroom or laboratory. Any learner may have just had an argument with friends or family and arrive to take the test in a distraught state. The possibility of these random psychological factors makes it all the more important that you do what you can to create a favorable testing environment.

Plan your test carefully. You will recall that the single most important characteristic of a good test is validity. A major factor affecting the validity of a test is how comprehensively it samples learners' knowledge. Careful planning can help you to develop tests that are comprehensive samples of your learners' knowledge.

The steps in planning a test are few and simple. First, you should identify the specific cognitive learner performance objectives for which you want to measure learner achievement. Second would be to list those objectives on a sheet of paper or planning form. This list of cognitive learner performance objectives should describe the content to be covered by your test.

Third, you should review your daily lesson plans or learning packages, if necessary, to identify additional, more specific content to be included in the test. Sometimes, learner performance objectives are not stated in enough detail to identify the specific content to be tested. Should you find that to be case, the greater amount of detail provided in your lesson plans or learning packages will be helpful.

Fourth, you will need to determine how many items to develop for each objective you have listed. Different objectives may well need to be covered by different numbers of items. How many items to include for a single objective depends on how important that objective is compared to the other objectives covered in the test.

Some indicators of this relative importance are (1) the amount of time spent on instruction for each objective, (2) the amount of emphasis placed on instruction for each objective, and (3) the amount of material covered by each objective.

There may be other indicators of relative importance to consider in your own situation. In any event, the more important or extensive an objective is, according to these indicators, the more items you should include to cover it. Furthermore, the relative importance of objectives should not come as a surprise to your learners.

In determining the number of items to include for each objective, you are determining the relative weight to be given to different content in your test. If you choose, you can carry your planning forward another logical step. You can determine the point value or other weight to be assigned to each item or cluster of items.
As you do so, it is very important to remember that point value should reflect the same relative importance of the objectives. If Objective A is covered by twice as many items as Objective B because it is twice as important, then the total point value of the items for A should be twice the point value of the items for B.

Sample 1 illustrates one planning form that you might find convenient to use in organizing your test planning efforts. You can list your cognitive learner performance objectives or other descriptors of content in the left-hand column. The number of items to be developed for each objective can be listed in the second column; the type of items (e.g., multiple-choice), in the third column; and the total point value for these items, in the right-hand column.

Sample 1

<table>
<thead>
<tr>
<th>Test Planning Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Learner Performance Objective</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>