CHAPTER 5

Assessing Learning Performances

CHAPTER OVERVIEW

After considering the content to be learned, instructional objectives, and the students to be taught, one important task remains before you can actually plan lessons. You need to design the assessments that will show what students have learned from your lessons. You may ask, "Why plan the assessments before instruction? Don't the tests come after the lessons?" The answer is based on the principle of "backward design"—start with the end in mind (Wiggins & McTighe, 1999). At this point in our planning process, we need to know not only where we are headed (the objectives for student learning), but also how we will assess those objectives. Only then can we design lessons that systematically develop the intended learning.

A well-constructed assessment provides critical information regarding the extent of students' learning and provides feedback regarding the strengths and weaknesses of instruction. Assessment involves:

a. determining what students already know and need to learn (preassessment);
b. using the information to make instructional planning decisions;

c. selecting assessment methods that align with content standards, benchmarks, and instructional objectives;
d. measuring how well students have attained these standards and objectives;
e. providing informative and useful feedback to students and parents; and
f. modifying instruction to better address student learning needs.

SECTION 1. EDUCATIONAL EVALUATION, ASSESSMENT, AND GRADING

Opening Activity

Read the following classroom dialogue between a teacher and a parent at Ken Cowan's school. Then answer the questions that follow.

MRS. CAMPISON (PARENT OF AN EIGHTH-GRADER): Mrs. Samson, my daughter, Loretta, is in your eighth-grade class this year. I'm afraid that she is very upset about the grade she received from you in math this marking period. She has complained often since she received her report card the day before yesterday.

MRS. SAMSON (EIGHTH-GRADE TEACHER): Mrs. Campison, I'm sorry about Loretta's reaction to the grade she received. My purpose in teaching is to assist students to be more confident and knowledgeable persons rather than to make them unhappy.

MRS. CAMPISON: Is it true that you are a harder grader than Mr. Carson? Loretta told me that students in your class do more work than her friends in Mr. Carson's classroom.

Practice Point

MRS. SAMSON: I won't compare my requirements to another teacher's requirements, but, it is true that I expect students to do their best work. I have found that if I have high expectations for my students, they will make more learning progress than if I were less demanding.

MRS. CAMPISON: But what about the students who are frightened by your approach to teaching, as Loretta is? What do you do to accommodate these students?

MRS. SAMSON: First, very few students do not benefit from my approach. Consider that we are at the end of the first marking period. Some students need more time to understand that I am serious. Second, I believe that it is the teacher's responsibility to motivate students and to provide the best instruction possible. However, it is also the teacher's responsibility to judge how well the students have learned what has been taught and to assign a grade based on that judgment.

I appreciate the concerns you express about Loretta's schoolwork and hope you will encourage her to respond positively to the challenge to do better in school.

MRS. CAMPISON: Where is the place for effort in your scheme of things? Loretta tells me that she tries as hard as she can but she still fails to satisfy you.

MRS. SAMSON: Effort, if it is strong and persistent, results in improvement. I am absolutely convinced of that.

MRS. CAMPISON: I see that we are not getting anywhere with this discussion. I do intend to speak with Loretta and to offer her additional assistance. I will also encourage her to pay more attention in class and to ask for help from you when she needs it. However, I must say that I find your approach to teaching children unduly threatening and unbending. After all, these children are only 12 and 13 years old. I believe they ought to enjoy their youth and be encouraged to cooperate with each other rather than to compete for your approval and for grades.

MRS. SAMSON: Mrs. Campison, if you and I support each other and motivate Loretta to do her best work in my classroom, I am confident that we can make a significant contribution to her educational development.

What is your reaction to the two persons featured in the dialogue? Do you find yourself more in agreement with the ideas expressed by Mrs. Campison or Mrs. Samson? Why?

Section 1 Objectives

After you have completed this section you will be able to:

1. explain the roles of assessment throughout the instructional cycle;
2. compare and contrast common educational terms such as preassessment (diagnosis), assessment, norm- and criterion-referenced evaluation, formative and summative evaluation, and grading; and
3. list common assessment procedures on a continuum from most controlled to most natural.

Educational Evaluation

Educational evaluation is a systematic process that leads to a judgment about the worth or merit exhibited by a person or persons or by an instructional program. Educational evaluation is considered systematic because it is planned, organized, and completed periodically, and its judgments are subject to revision or redress through established review procedures. Compare this systematic process to the evaluations we make in everyday life. How often do we decide whether something is attractive or unattractive, fair or unfair, good or bad? Normally we are not called upon to defend our evaluative decisions. However, the evaluations that teachers make are fair game for public scrutiny.

Often an educational evaluation entails judgments about both the learner and the instructional program that produced the learning. The conversation between Mrs. Campison and Mrs. Samson involved the evaluation of a student's performance. However, the parent implied that the teacher's instructional approach was a factor affecting the student's performance. Parents have the right to demand that teachers can defend their educational evaluations with appropriate reasoning and evidence using language that a noneducator will understand.

Educational evaluations can be made at any time and for a variety of purposes. Two of these purposes are labeled formative and summative. **Formative evaluation** provides information for improvement while the person or program has the opportunity to improve. Thus, formative evaluation helps a teacher "form" better decisions about student learning and instructional success. Examples of procedures used for formative evaluation of learners are pretests, self-tests, quizzes, drafts of assignments that are reviewed and returned for revision, and practice exercises. Examples of formative evaluation of instruction include exercises to check for understanding and short questionnaires in which students are asked to identify what objectives they have and have not achieved as well as reactions to the classroom activities.

Summative evaluation is used to make educational decisions about persons or programs after instruction terminates. Examples of procedures used for summative evaluation of learners include authentic projects, unit tests, final exams, term papers, student teaching evaluations, report cards, decisions to retain or promote a student, and the like. Examples of summative evaluation of instruction are student evaluations of courses and teachers, formal evaluations of teachers by supervisors, tenure reviews, and other procedures.

The purpose of the evaluation determines whether it is formative or summative. In reality, a given evaluation procedure may serve both purposes. Think for a moment about a unit test that is designed as a summative evaluation of students' performance after the completion of one unit and that also serves a formative purpose to guide students toward improved performance in the next unit.

Educational Assessment

If educational evaluation is a decision about worth or merit, how do we define terms such as assessment and grading? *Assessment* is the process of measuring the quantity and/or quality of a behavior or the indicator of that behavior. It is the foundation information upon which an educational evaluation decision is based. When teachers assess, they determine what students know about some content or their level of skill in performing some task. Teachers might also assess students' motivation to engage in further study of that content.

Assessments can be placed on a continuum between controlled and natural. A highly **controlled assessment** suggests the use of a special testing environment—perhaps having chairs arranged to discourage copying from neighbors, no books or papers on desks, or the requirement that the test be completed within a given time period. In a controlled assessment, the learners are always aware they are being assessed. The most common form of controlled assessment is a standardized or teacher-made paper-and-pencil test, typically administered in the classroom. Teachers can choose other kinds of controlled assessments when appropriate. For example, the observation of a student product or presentation offers a different view of student skill or achievement than does a paper-and-pencil test. Less controlled assessments include written homework assignments, interviews, questionnaires and checklists, observations, and examination of existing records.

Most Controlled ←				→ Most Natural
Objective and essay tests	Student presentation in classroom, products created in classroom	Observations, interviews, questionnaires completed in a controlled environment—e.g., classroom, testing center	Out-of-school assignments (products, written work, logs, and diaries)	Observations in natural settings, e.g., bus stop, library; examination of existing records

Figure 5.1 Assessment Continuum: Most Controlled to Most Natural
From *Teaching as Decision Making: Successful Practices for the Elementary Teacher* (3rd Ed.) (p. 127), by A. J. Starko, et al., 2003, Upper Saddle River, NJ: Merrill/Prentice Hall. Copyright 2003 by Pearson Education, Inc. Reprinted with permission.

Natural assessment procedures are at the far end of the spectrum. Natural assessment requires no artificially constructed testing environment, and the learners are not aware that they are being assessed. For instance, the teacher might observe students while in gym or the library or check the kind and number of library books students choose. A continuum of assessment procedures from most controlled to most natural is shown in Figure 5.1.

The Relationship Between Instruction and Assessment

Consider the following scenario:

A first-year teacher working with Ken Cowan on the Zimbabwe unit has begun assessing her eighth-grade students' first drafts of a formal research report. Using a minimum of three resources, students created two-page reports on the differing perspectives of White farmers and landless Blacks. The teacher is shocked and dismayed by the poor quality of student work. Only 4 of the 25 students have followed the written instructions accurately. Perplexed by the number of students who have submitted papers without important headings, specific comparisons between White farmer and landless Black points of view, and properly labeled timelines, the teacher reflects upon possible reasons for such results.

While some teachers might be inclined to find fault with the students by concluding that they lack research experience or ability, this teacher recognizes that the sheer number of students who failed to meet expectations indicates other causes. An honest assessment of her instruction yields some important insights: the need for clearer expectations and new forms of tiered instruction to promote success. Perhaps a minilesson on comparing and contrasting ideas using a Venn diagram will assist students in their effort to analyze the contrasting perspectives. The teacher considers creating a checklist of report components that students will use as a self-assessment before turning in their final report. Finally, she weighs the potential benefits and drawbacks of presenting students with a model report she could create herself.

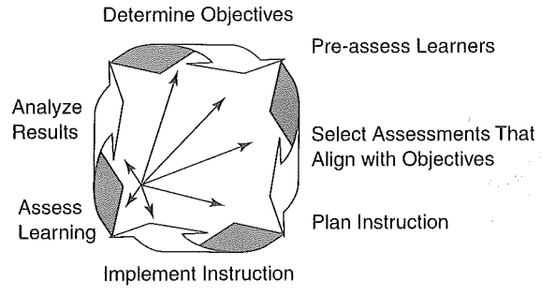

Figure 5.2 The Instructional Cycle
From *Teaching as Decision Making: Successful Practices for the Elementary Teacher* (3rd Ed.) (p. 128), by A. J. Starko, et al., 2003, Upper Saddle River, NJ: Merrill/Prentice Hall. Copyright 2003 by Pearson Education, Inc. Reprinted with permission.

This first-year teacher has made an important connection between assessment and instruction. Her students' work provides critical information for the improvement of her own instruction and for future planning. Figure 5.2 shows how selecting instructional objectives, preassessing learners, planning instruction, assessing learning performances, and analyzing results are all part of a continuous cyclical process.

When viewed as an integral part of the instructional cycle, assessment data can be used to generate a wealth of information regarding the quality of our goals and objectives, the effectiveness of our instruction, and of course, the strengths and needs of our students.

Teachers currently use assessment methods at each phase in the instructional cycle (represented by the inner arrows in Figure 5.2). We call this *continuous assessment.* For example, before designing lessons, teachers preassess students to see who may have already learned the objectives or who is lacking prerequisite knowledge. During the lessons themselves teachers informally assess students through the use of observation, assigned work, discussion questions, and one-to-one conferences. At the end of a series of lessons, teachers assess their students' learning of the most important standards and objectives.

Why have teachers begun to embed such formative assessments throughout the instructional cycle? Teachers use the information to clarify and revise their expectations, objectives, and instruction to better address student needs. As a consequence, greater numbers of learners will be successful, and all students will be sufficiently challenged.

One method for continuous assessment is the analysis of student work samples. For example, every week one teacher selects five different students from each period whose work (e.g., class work, homework assignments) will be analyzed in detail. This teacher becomes genuinely curious about why these five students perform as they do, and she attempts to discover the best way to help each one. As she analyzes the work samples, she consults other resources—teachers, books, experts or other professionals—to help her understand the learners' needs more deeply.

Working together with other professionals to interpret and understand student work can be a particularly powerful strategy. In one study of this process (Langer, Colton, & Goff, 2003) teachers changed from seeing student work as merely a grade (or check) in the grade book to analyzing the exact nature of a student's learning and using that information to make instructional decisions. Consequently, their students learned more.

Preassessment

A key idea in this book is that teachers continuously gather information about student learning and use that information to increase student success. One of the more important decisions is where to begin instruction so it isn't too hard or too easy. A powerful tool for making such decisions is preassessment, or diagnosis. For example, if Ken Cowan is about to teach a lesson on the differences between a traditional life in Zimbabwe and a modern life, he should think about two kinds of assessment: *preassessment* and *postassessment.*

When teachers preassess student performance prior to instruction, it is to determine what the student knows or is able to do in relation to the lesson or unit objectives. Teachers also may determine whether students' feelings are positive, neutral, or negative toward the content. They can then use this information to tailor the instruction so it is at the appropriate level of difficulty and interest.

When the teacher assesses student performance after instruction (postassessment), the concern shifts to determine what the students have learned as a result of the lesson and activities. One way for teachers to know if their instruction has had desired results is to compare the students' performance on the preassessment and postassessment of the same objective.

Preassessing students' entry-level competence requires two decisions: what methods to use to gather the information and what content and skills to assess. Preassessment can be accomplished through both *formal* and *informal methods.* An example of a formal method would be a written pretest, similar in style, length, and degree of difficulty to the test that you will administer when the lesson or unit instruction has ended. Formal methods have the disadvantage of feeling like a test—long and possibly scary. One strategy some teachers use is to spread the parts of the pre-test over two or three days, so it does not overwhelm students. Figure 5.3 shows preassessment purposes and methods.

In general, however, you will want to use an informal method, spending as little classroom time as necessary to secure the required information. For example, Ken

Preassessment Purposes and Strategies

Area Assessed	Examples of Preassessment Strategies
Interests/feelings	Survey, questionnaire, free-write, rating scales, KWL
General learning skills	Perform (e.g., use equipment, write, do math)
Prior experiences	Free-write, log, daily journal, discussion
Knowledge of content	Pretest, anticipation guide, game, perform task(s), self-assessments (KWL), concept maps

Figure 5.3 Four Categories of Preassessment
From *Teaching as Decision Making: Successful Practices for the Elementary Teacher* (3rd Ed.) (p. 131), by A. J. Starko, et al., 2003, Upper Saddle River, NJ: Merrill/Prentice Hall. Copyright 2003 by Pearson Education, Inc. Reprinted with permission.

Cowan could ask students to list up to three characteristics of a traditional life and three characteristics of a modern life and then choose which life they would prefer to lead. He would collect the papers and analyze each student's level of understanding to determine who may need more or less challenge.

In regard to the content of a preassessment, think of it in two different ways. First, you need to assess the general knowledge and skills you hope students already possess. What *prerequisite* knowledge and skills do students need to approach the objectives in this unit? Do students have the necessary background knowledge to read successfully the written material you have gathered? Can they take notes on lectures and reading? What do they know about library research? Can they do the math required to do the graphing? How well can they cooperate with others during group work? The easiest way to preassess these general skills is to ask students to use them in a simple setting. For example, imagine Mrs. Samson wants to know how well her students can create graphs to represent data they collect about Zimbabwe. She might give them a small set of numbers, explain the task, and see what she gets. It is likely some students will need explicit instruction on this skill while others may not. She may form project groups that include both kinds of students.

Second, preassessment will help you determine each student's prior learning about the content to be taught. The unit concept map and objectives (see Chapters 3 and 4) are very helpful when designing a preassessment of unit content. After completing his concept map Ken Cowan would know that in Zimbabwe a traditional life is typically based on agriculture; and is rural, village-based, family-oriented, and slowly changing; whereas a modern life tends to be urban, technological, impersonal, fast-paced, and dynamic. His preassessment would ask students to list ideas they have about traditional and modern life in Africa or Zimbabwe. Then he would look for the ideas listed above in the students' responses.

This content-specific preassessment may show that some students already have learned what he planned to teach. He will plan independent study or special projects for these students. Such preassessment is extremely useful for teaching decisions. You can use your analysis of the responses to adjust the complexity of the content to be taught, the nature of the activities you will use to initiate study of the unit, and other planned activities and assessments. Further, you will be able to modify (or differentiate) your instruction so individual students are sufficiently challenged and supported.

REFLECTING ON THE IDEAS

If you were Ken Cowan, what would you wish to know about the eighth-grade students' entry level competence before you begin teaching the lesson on the differences between a traditional life in Zimbabwe and a modern life?

You might have considered: Are the students able to read a chapter and gain meaning? Can they take notes from a lecture or movie or field trip? How much do they know about the process of comparing and contrasting?

Grading

How do the terms *assessment* and *evaluation* relate to the issue of *grading?* Ken Cowan has used a preassessment to evaluate how well his students were prepared to study the unit on Zimbabwe. When the results were gathered and examined, he made a judgment (evaluation) about whether the students were ready to study the unit as tentatively planned or if adjustments needed to be made. At the end of the unit, he will postassess the students, evaluate what they have learned, and will assign a *grade* to communicate his judgment to the student, parents, and interested others. Bear in mind that a grade is a label that our society uses to represent value earned. It is a kind of certificate that communicates earned achievement to educators, employers, and the general public. Remember that the value of a grade is only as good as the assessments upon which the grade is based. If they are to be useful to society, grades have to be sufficiently accurate so that they can predict future performance. Thus a student who receives nothing but A's in middle school mathematics should expect to receive similar grades in senior high school.

No doubt you are aware of the intense debate about the relative worth of the grading process. The decision whether to assign grades, the process used to determine the value represented by the grades, and the particular symbols used as grades are political decisions as much as educational decisions. Some educators argue that the symbols traditionally used for grading do not clearly represent the complexities of student achievement and that more specific information would be helpful (Marzano, 2000).

Norm-Referenced and Criterion-Referenced Evaluation

A classroom is sometimes viewed as containing a group of 20 to 35 students who compete against classmates for grades, awards, scholarships, and admission to college. College students may be accustomed to seeing an instructor put a grade distribution curve on the chalkboard showing how many students received an A, B, C, D, or F. Although the concept of normal distribution may remain mysterious, students know that the instructor arranged all the grades in a row and then established cutoff scores between grades—for instance, 92–100 = A, 85–91 = B, and so on.

This view of grading is referred to as **norm-referenced.** The average score and other scores that deviate from that norm become critical determinants of an individual's grade. Scores that fall well above the average get A's and B's; scores that fall well below the average get D's and F's. Scores that cluster around the average get C's.

Norm-referenced evaluation has one major disadvantage. As student performance improves, the average is higher, and the teacher is likely to raise the curve to retain a similar proportion of A's, B's, and C's. Thus, students experience the phenomenon of working harder or doing better without seeing any grade change. A second major disadvantage is that the value of the grade is difficult to interpret from one group to the next. Since the grade is determined by the performance of students in a particular group, a grade in one eighth-grade science classroom taught by teacher A may represent much more or much less educational value than an identical grade in an eighth-grade classroom taught by teacher B. Large testing programs such as the National Assessment of Educational Progress (NAEP), Scholastic Aptitude Test (SAT), and American College Testing (ACT) can avoid this problem by comparing an individual's score against the average derived from a national sample group.

In contrast, **criterion-referenced grading** considers each individual's performance against predetermined performance standards (or criteria). An evaluation that is criterion-referenced enables the teacher to determine with confidence whether a student's response reaches a predetermined level of accomplishment. Using standards as the definition of success rather than student-to-student comparison makes it theoretically possible for all students to be successful and receive A's. The list of standards (or criteria) that define the desired level of proficiency is referred to as a *rubric* (Wiggins, 1996; 1998).

Some assessment models and types are more appropriate for criterion-referenced grading than are others. One example of a criterion-based teaching/learning model is called **contract learning.** In such a model, the standards for receiving a given grade are described to all students. Each student contracts with the teacher to perform certain tasks at a given quality level in order to receive the agreed-on grade. Objectives and activities are identified, criteria determined, and deadlines established. Contract learning has the advantage of permitting a student to concentrate on certain subject areas or units within a subject and not on others. It also helps prevent the teacher from labeling students, for instance, as A or C students, since students can choose the grade they will seek to achieve.

An example of contract learning is the Brain-Flex program implemented in a secondary school in New South Wales, Australia (Bounds & Harrison, 1997). Each student in the program completes two or three independent projects during a school year. These projects enable them to flex their minds in pursuit of intellectual topics and outcomes of interest to them rather than those selected by the state. The topic must have sufficient breadth and depth to justify the effort that will be devoted to it. That decision is made after the student has written an explanation of the topic and justified its worthiness. Once a suitable topic has been determined, a learning contract is developed with the student's tutor (teacher). The contract includes a rationale of what the student expects to learn, clearly written goals and objectives to be achieved, and standards for assessment of the expected outcomes. Each student is free to use whatever facilities, equipment, and materials are needed to study the topic. Faculty and staff are available to provide counsel and assistance as needed. At the end of the contract, students submit the outcomes of their study and a written summary of what they have learned.

Another example of criterion-referenced grading is known as **mastery learning** (Carroll, 1963; Bloom, 1984). Mastery learning is based on the belief that all students can be successful in achieving all objectives if additional learning time is allowed for those who need it. Mastery learning, if it is to be successful, requires that teachers use clearly stated objectives, employ preassessment and other kinds of formative assessment procedures, and adjust instruction using a variety of strategies and learning activities. Finally, they must be prepared to reteach students when necessary, using alternative methods and materials. Advocates claim that if these assumptions are met and if a sufficient time for learning is provided, all students can master all objectives.

In all these examples criterion-referenced evaluation can be applied to the grading of either tests or authentic performances. Remember it is not the form of the assignment

that makes it criterion-referenced; it is how it is scored—by comparing student performance to a set of criteria that defines "mastery."

Tomlinson (2001) suggests that grading needs to accommodate individual differences and variations in instruction. For example, how would a teacher grade a student who is making excellent progress but is working at an instructional level below the other students? Similar puzzles arise for a student who is struggling but working on exceptionally advanced material. In most school settings, these decisions are only partially under the control of the teacher.

SECTION 2. DESIGNING CLASSROOM ASSESSMENTS

Section 2 Objectives

After you have completed this section, you will be able to:

1. explain the importance of validity when evaluating student learning;
2. describe strategies for increasing the validity and reliability of classroom assessments;
3. apply recommended guidelines to construct objective and essay tests;
4. explain how authentic performance assessments can aid evaluation decisions; and
5. design performance tasks to assess student learning.

Classroom assessments used for grading may come from published curriculum materials (e.g., the teacher's guide for a textbook) or may be designed "from scratch" by the teacher. It is rare, though, that a textbook assessment is well matched to state standards, particular students, and what is actually taught. Therefore, most teachers design their own assessments or at least adapt published ones for a better fit. This section introduces key concepts and methods to help you design high-quality assessments. The design and interpretation of assessment information is an extremely complex process. Our presentation is just a beginning; ideally, you will have the opportunity to study this topic in greater depth.

Validity

Validity is the property of being relevant, meaningful, or effective (Merriam-Webster's, 1993). *Educational validity* can be defined as the property of an assessment that makes it an accurate measure of what it purports to measure. Although there are several kinds of validity, only content validity is directly under the teacher's control and responsibility. *Content validity* has often been described through the following teaching prescription:

> Tell them what they will learn, teach them what you told them, and assess what you taught.

However, achieving a high level of content validity is more complex than this simple prescription suggests. Fortunately, your analysis of the content you will teach and the use of clearly stated objectives are the foundations on which valid tests are developed. Two dimensions of content validity are most important. In the first dimension a valid assessment is one that provides an adequate coverage of the content taught in the unit.

In the case of the Zimbabwe unit, adequate assessment of the content includes all the key objectives, including the concepts and generalizations that organize the unit.

If some aspect of the content is missing from the assessment, students can rightfully argue that too much emphasis was placed on some objectives and too little on others. Another common violation of validity occurs when students are assessed on skills or knowledge not taught in class but acquired previously—material learned in earlier courses or through students' personal experiences. For example, consider the injustice that is created when a social studies teacher permits a student to submit a unit project on colonial America that was the result of a summer family excursion through the original 13 colonies. Obviously, one family provided a student with opportunities for cultural enrichment that were far beyond the capacity of other families. To give a substantial grade advantage to the student for this project compounds the effects of economic and class inequity.

The second dimension extends the issue of validity into the learning process. It is necessary not only to "assess what you taught" but also to assess at the same level of learning specified in your objectives. Perhaps you know of instructors who promise lofty aims such as enhancing thinking skills, learning to make informed judgments, and exploring ways to solve problems and then test for the recall of bits and pieces of information. To avoid this error in assessing the Zimbabwe unit, adequate coverage must include (1) questions on the application of concepts such as traditional life and modern society, racism, and colonialism, and (2) tasks that require students to analyze data and to create original products such as a site map and a tribal mask.

Reliability

A second important evaluation concept is reliability. Whereas validity is concerned with accuracy of assessment, **reliability** is concerned with consistency (or stability) from one performance to the next. Educational assessments that have high levels of reliability ensure that if a student were assessed again on the same test, he would achieve an identical or nearly identical score. Thus, if an assessment is both valid and reliable, the teacher can rely on the information as being a reasonable indication of what the student has learned. Validity assures that the assessment measures what you think it measures. Without validity it makes no difference whether the assessment is reliable or consistent. These two concepts can be applied to both written tests and performance assessments.

Enhancing the Reliability and Validity of Teacher-Developed Written Tests

Teachers develop a wide variety of written tests, which can be classified into two types: objective and constructed response. In an *objective test* students are expected to identify or provide the right or best answer. A multiple-choice test is the most common example of an objective test. Other examples include matching, true/false, and fill-in-the-blank. In a *constructed response test* students are expected to build a response requiring the planning, organization, and display of multiple elements. Essay tests and reports are the most common examples of a constructed response test. The following section is designed to help you to plan, design, and grade objective and essay tests.