CHAPTER 5

Assessment of Student Learning

Rena B. Lewis, Ph.D.
Professor of Special Education, San Diego State University
Faculty member, San Diego State University–Claremont
Graduate School Doctoral Program

Editor's Notes:
Rena Lewis's chapter examines specifically the assessment component of the multicultural course change model presented in Chapter 2. She begins by engaging readers in one of the assessment strategies she recommends. The author maintains our engagement throughout the chapter with her highly accessible wisdom communicated through personal examples and activities. After taking us through an analysis of our current assessment practices, she offers two approaches to achieving more equitable evaluation of student progress in a course. The first approach is to modify traditional assessment tools. The second involves employing alternative strategies because "no course can be considered multiculturally infused if it utilizes only traditional assessment procedures." Lewis' suggestions focus on the multicultural goal of improving diverse students' learning through assessment strategies sensitive to a range of learning needs. However, as she also emphasizes, moving assessment from exclusive to inclusive to transformed supports the knowledge acquisition of *all* students. ■

In college courses, the process of assessing student learning does not begin on the day of the first exam. It begins far earlier: when the course is first conceptualized, when course goals are identified, and when the course syllabus is prepared. At these stages, the professor determines the content of the course, the instructional strategies that will be used to teach that content, and the types of activities in which students will participate to demonstrate what they have learned.

In this chapter, *assessment* is defined as the process of gathering information about student performance in order to evaluate progress toward mastery of course goals. This process involves a series of steps including specification of desired student outcomes, selection of an assessment strategy, design of the assessment task, data collection, and evaluation of student performance.

In recent years, leaders in the higher education community have begun to focus increased attention on the ways in which student achievement is assessed in an attempt to improve the evaluation process (Banta, 1992; Ewell, 1991). One major strategy has been the development of new types

of assessment tools as supplements to or replacements for traditional measures. Before reading further in this chapter, please experiment with one of these tools, the concept map. Take a blank sheet of paper and write the words "classroom assessment" in the middle. Draw a circle around those words, then jot down other concepts that come to mind when you think about classroom assessment. Draw lines to join related concepts together as needed. For example, if I were drawing a concept map of "desserts," I might begin with subordinate concepts such as "sweets" and "fruit," then develop each of those (e.g., the concept "sweets" might be linked with "cakes," "pies," "cookies," and so on). Your map for "desserts" might look quite different, however. Concept maps are used for several purposes, including gathering pretest data on learners' knowledge of and attitudes toward a subject prior to instruction. For example, in a women's studies course, students were asked to prepare a concept map of "feminism" on the first day of class; the instructor then used the results as a basis for class discussions and to structure lectures and other classroom activities (Angelo & Cross, 1993).

In developing your concept map on "classroom assessment," you likely included one or more of the most typical tasks used to assess student learning: exams, written assignments, and projects. Some college courses, particularly those with large numbers of students, require only exams (for instance, a midterm and a final). Others combine exams with a project such as a term paper. In others, frequent quizzes replace exams and students complete several written assignments in lieu of a major term project.

However, traditional assessment tools are not always the best choice for nontraditional students. Included in this group are students from diverse cultural and linguistic backgrounds, women (particularly in majors where women are traditionally underrepresented), and students with disabilities. Traditional assessments present three barriers to the optimal performance of students such as these. First, there may be a mismatch between the content knowledge that students bring to the class and the knowledge required to successfully participate in assessment activities. For example, one woman student nearly failed a midterm exam in a statistics and probability class, not because of her lack of knowledge of the course content but because many of the test questions asked students to compute the probability of different types of poker hands, a game with which she was not familiar. Second, there may be a mismatch between students' skills and the skills required to demonstrate learning on traditional assessment tools. For example, a multiple-choice exam written in dense, cryptic English may not fairly assess how much a student from Cambodia knows about the United States Constitution. Likewise, an in-class writing exercise may unjustly penalize the student with a physical disability who writes well, but slowly.

Third, there may be a mismatch between the values held by students and those inherent in the assessment activity. For example, in one of my earliest teaching experiences, I tried to motivate Native American students from the Navajo Nation to excel on the first exam of the course by promising public recognition of all who earned A's. One of my students kindly took me aside and explained that in his value system the achievement of the group is prized over the achievement of an individual.

It is critical to note that *all* students, not only those from nontraditional backgrounds, may benefit from efforts to rethink our classroom assessment procedures. Students in every group, including White middle-class males, exhibit a range of individual differences in how they best demonstrate what they know.

There are two major approaches to overcoming barriers that assessment procedures may place on students' abilities to demonstrate what they have learned. In the first, traditional assessment tools are modified to make them more equitable for nontraditional students (and more accessible to all students). In the second, alternative strategies for assessment are used instead of, or in addition to, traditional methods. This chapter will discuss both approaches. The chapter begins by exploring the factors to consider when selecting or designing assessment tools. The next sections provide a strategy for analyzing and then modifying assessment practices. The final section presents alternative methods of assessment and principles for their incorporation into a course curriculum.

The suggestions offered in this chapter are based upon a set of assumptions regarding the assessment of student learning. These assumptions serve as the guiding principles for the design of appropriate assessment procedures for all students.

- *The tasks selected for assessment must be directly related to course goals.* This is a validity issue. If one course goal is that students will be able to deliver a 5-minute extemporaneous speech, a written exam on extemporaneous speaking is not a valid measure (McLoughlin & Lewis, 1994).
- *Assessment tasks should reflect the most important course outcomes.* What the task reflects is also a validity issue. It is usually not possible to assess all of the outcomes that a course is designed to bring about. When choices are made, priority should be given to major topics, not minutia. Also, the topics selected should be a representative sample of the entire domain under assessment, not just one or two aspects of it. If a course covers the history of the United States from the colonial period to modern times, half of the final exam should not be devoted to questions on the Civil War.
- *Assessment techniques should be designed to bring out the best performance of students.* Implementing this principle requires that the professor consider student strengths when making decisions about assessment strategies. Although this consideration is important for all students, it becomes essential when attempting to promote the academic success of students from nontraditional groups.
- *Assessment strategies must not only be effective, they must be efficient.* Holding 30-minute conferences with each student at the end of the term may be an effective way to gauge learning. It is not efficient when there are 200 students—and only one instructor—in the class.

FACTORS TO CONSIDER IN ASSESSMENT

When professors sit at their desks, pen or keyboard in hand, and draft a course syllabus and course requirements, they are making several decisions about assessment. These decisions may be made on the basis of history and tradition, that is, the professor's own experiences as a student and member of an academic discipline ("in my field, we only use essay exams") and/or on the basis of past teaching experiences ("This is how I've always done it and it works"). Or professors may decide to try a new technique that they have read about in the literature or heard discussed by colleagues. In many cases, however, decisions are made without conscious consideration of the factors that make up the assessment process. Those factors become important when thinking about ways to improve assessment prac-

tices, and it is useful to have a model that describes them. These factors include (1) the characteristics of the assessment task, (2) how tasks are weighted, (3) the scheduling of assessments, and (4) procedures for grading and feedback to students.

The Assessment Task

Exams, term papers, oral reports, tests, abstracts, essays, reaction papers, debates, critiques, quizzes—these are but a sampling of the different types of assessment strategies that professors use to obtain information about student learning. Each strategy also represents a range of options; exams, for example, can be objective or essay, in-class or take-home, open or closed book, performance or paper-and-pencil, and so on. One way to sort through the options when selecting or designing a specific assessment task for a particular course is through analysis of the dimensions that characterize all assessment tasks.

Task Level

The level of an assessment task refers to the type(s) of cognitive demands it places upon students. Asking students to name the current head of state in country X clearly imposes different demands than asking them to discuss the forces that led to that leader's rise to power. One way to conceptualize task level is through the *Taxonomy of Educational Objectives* proposed by Bloom and his colleagues. In the cognitive domain (Bloom, 1956), this schema classifies course goals into six hierarchical levels:

- *Knowledge,* or the recall of information
- *Comprehension,* or the understanding and interpretation of information
- *Application,* or the use of information
- *Analysis,* or the breakdown of information into its component parts
- *Synthesis,* or the production of new information
- *Evaluation,* or judging the value of something in relation to a specific purpose

Although this schema does not apply equally well to all disciplines and all courses, it provides a basis for thinking about task level and the cognitive demands that different types of course goals impose.

Consider, for example, two assessment tasks that might be used in a beginning level foreign language class. One goal of the course is that students will learn basic vocabulary. This is a knowledge goal (as defined above) that could be assessed with a paper-and-pencil test. A second goal is that students will be able to use their new knowledge of the language in authentic communication situations. This is an application goal. To assess it, the instructor might require students to interview a native speaker, then present an oral report of that interview to the class.

Input Mode

This dimension is concerned with the types of inputs students receive prior to the assessment task. The most obvious type is the set of instructional activities that make up the course: lectures, text, and other required readings, demonstrations, discussions, and laboratory work. In some cases, course activities are supplemented with additional experiences that form part of the assessment task; the interview of the native speaker, discussed below, is an example.

The directions for the assessment task itself are another type of input. Directions may be as brief as one sentence written at the top of a test ("Read each question carefully; answer all questions"). If the assessment task is a writing assignment such as a term paper, the professor may supplement

written instructions with a thorough discussion of the parameters of the assignment and grading criteria.

The knowledge, skills, and experiences that students bring with them to a course represent a third type of input. Often called "prior knowledge," this input is the foundation upon which students build their competence in specific course content. Professors base their expectations for student performance upon assumptions about prior knowledge ("Of course sophomores know how to use the library"); these assumptions form part of the hidden curriculum. For example, in most college classes, professors assume that students can read the textbook and other required readings with understanding, that they are able to take notes from lectures, and that they know how to study for and take exams.

Response Mode

Response mode is the medium students use to express their ideas when completing an assessment task. The most typical response modes in college courses are writing and speaking. Students write exams, reports, projects, and papers; they present oral reports and participate in class discussions and debates. Courses related to the arts may require other response modes: singing, playing an instrument, designing a stage set, drawing, painting, sculpting, and so on. Each response mode can be further described in terms of the type of product required. For example, in writing, the response may be as simple as writing one letter to answer a multiple-choice test question or as complex as writing an original poem, short story, or play.

Use of Aids

Aids are supplements that assist students as they respond to assessment tasks. For instance, students might be allowed to use a calculator on a math exam, a print, or electronic dictionary when writing an essay exam. In an "open book" exam, students have access to their textbook(s) and sometimes class notes and handouts. Likewise, some professors encourage (or require) students to take advantage of aids as they complete written assignments. For example, when writing a term paper, students might take part in a library tour, attend a term-paper clinic, use a word processor with a spelling and grammar checker, review a model paper from a previous semester, and confer with the course instructor or a teaching assistant.

Language Requirements

Assessment tasks differ in the types of language demands they impose on students. In college classes, however, tasks are typically language-intensive, requiring high levels of skill in reading and writing. When written language skills are de-emphasized, oral language skills often take their place. For some students, the difficulty level of the assessment task is directly related to the degree of language proficiency required. This rule holds true no matter what the subject matter of the course. Reading a chemistry textbook, giving an oral report in a marketing class, and writing an essay in a comparative literature course are all language-based activities.

Speed and Quantity Requirements

Speed and quantity requirements are concerned with how much work the assessment task entails and how quickly that work must be accomplished. These two variables can sometimes interact. A 25-item short-answer quiz does not appear lengthy if students have the entire 50-minute class period to complete it. However, the magnitude of the task seems to increase when the time period for the quiz is shortened to 5 minutes.

Participants

In many college-level assessment tasks, students work as individuals. Each student completes his or her exam; each student writes his or her own paper. In some situations, it is also possible for students to work in pairs or in small groups or teams. Group assignments can be structured in two ways: all of the members of the group can collaborate on all aspects of the task, or each student can take responsibility for a separate portion of the work. Another option is for students to first complete their work individually, then act as peer reviewers or editors for each other.

Location

Students can complete assessment tasks in class or at some other location. Although in-class exams are most common, some professors also use take-home exams. Other possible locations are university facilities (e.g., library or computer lab assignments) and community venues. For example, students might visit an art museum or see a play, observe a city council meeting, or complete a practicum in a business or public school setting.

Weighting

Another important factor to consider is the weighting of assessment tasks in relation to students' final course grades. Most professors use some type of system that assigns differential weights to various tests and assignments. Class quizzes might be worth 20 percent of the final grade, a library research paper 20 percent, the midterm exam 25 percent, the final 30 percent, and class participation 5 percent. In a class where the grade is based on three exams, each might make up 33 percent of the total grade; or, the first test might count for 20% and the other two for 40 percent each. The advantage to the latter arrangement is that students can use the first test to become familiar with the professor's style. In some courses, there is only one assessment task: a senior paper, a thesis, or some other type of major project. In cases like these, students can be graded as they complete portions of the task; typically, however, the final product accounts for a major part of the grade.

Scheduling

Scheduling is concerned with the frequency of assessment tasks and their timing during the semester or quarter. If quizzes are one of the methods used to evaluate student performance, they could be given every class session, every week, every month, or at varying intervals throughout the term (as is often the case with "pop" quizzes). Many professors structure their courses so that they are able to monitor student progress several times, including at least one assessment task early in the semester. When university calendars permit, the first test or assignment can be scheduled so that it can be returned to students before the official course drop date. Early assessment provides useful feedback both to the students and to the professor; because the course has just begun, there still is time to make changes. Scheduling assessment tasks at spaced intervals throughout the semester also has advantages. Students receive periodic feedback, and the bulk of their grade does not rest on tasks to be completed in a flurry of activity in the last weeks of the term. Professors also avoid some of the stress associated with the end-of-the-term grading crunch.

Grading

There are several variables related to the process of assigning grades to student tests and assignments. First of all are the criteria

used to judge student performance. Criteria can be objective, subjective, or somewhere in between. The answer key for a multiple-choice exam is an objective standard; a checklist or rating scale used to evaluate a written assignment is more subjective. In some courses, grading criteria are explicit; at the start of the course or as each new task is assigned, students are informed about the standards that will be used to evaluate their performance.

A second variable is who does the grading. In most instances, the grader is either the professor in charge of the course or a teaching assistant (although, with some assessment tasks, it is possible for students to participate by completing self-evaluations). When the professor delegates grading responsibilities to someone else, she or he must also communicate the grading criteria and, if necessary, train the assistant.

Third, once the grader has marked the errors on a test or made comments on a written assignment, a final grade must be assigned. Some professors use an absolute grading system. If there are 10 items on a quiz and the student misses 3, his or her grade is 7 or 70 percent. Others prefer to grade on a curve. In this system, students whose scores fall in the middle range of the class receive a C, those in the next highest range receive a B, those with the highest scores receive an A, and so on. The ranges associated with the various letter grades can be set statistically (e.g., those falling within one standard deviation of the class mean earn a C) or by "eyeballing" the data. The advantage to the use of curves is that some students will receive A's, even if the highest grade in the class is 65 percent. The disadvantages are that students compete against each other (rather than striving to meet course goals) and that minor variations in performance (e.g., 34 versus 37 points out of a possible 60) can result in different grades (C versus B).

A fourth consideration are the consequences for poor performance. Some professors offer students strategies for remedying poor grades; others do not. One strategy is a "forgiveness" policy; for example, the lowest quiz grade is dropped or the lowest test score is raised one letter grade. Or, students may be allowed to take a make-up test or complete an alternative assignment to offset a failing grade. The grades of the two attempts can be averaged, or the second grade can replace the first. Extra credit work is another option. In this system, students take on an extra project or assignment to raise their overall class average or as a hedge against a future low grade.

Feedback to Students

The assessment process is not complete until students have received feedback about their performance. Knowledge of results is one type of feedback, and it can be as simple as a grade on a test. Informing students about which test items they missed, their overall score, and their standing in relation to others in the class helps them evaluate their progress. Such information may or may not assist them to learn the material they missed or to do better the next time. Improvement is more likely to occur if corrective (or informative) feedback is given. This type of feedback includes information on the types of errors students have made and provides direction for rectifying those errors. For instance, if a student knows that his or her essay on the abortion debate received a grade of C– because its organization was difficult to follow and it failed to address both sides of the issue, he or she can use that feedback to begin to improve his or her writing skills.

When feedback occurs is another important concern, and a distinction from the field of evaluation (Worthen & Sanders, 1987) is useful here. In summative evalua-

tion, the purpose is to judge the quality of an undertaking *after* it has been completed. In formative evaluation, data are gathered *during* the undertaking in order to improve it. In assessment, tests and exams are typically summative strategies. Providing students with feedback on a draft of a term paper would be a formative strategy.

A final issue is the timeliness of feedback. For feedback to have an impact on student learning, it must be received in time for students to act upon it. Returning the midterm three weeks before the final gives students little chance to revamp their study regimen.

THE FIRST STEP TOWARD INFUSION: ANALYZING CURRENT ASSESSMENT PRACTICES

The first step in improving assessment for students from diverse groups is analysis of current practices. The factors described in the previous paragraphs provide a schema for analyzing current practices. As Table 5–1 illustrates, assessment strategies can be described in terms of the course goals they address, the assessment tasks themselves, and other dimensions such as weighting, scheduling, grading, and feedback. Analyzed here are the two strategies mentioned earlier for measuring student progress in a beginning-level foreign language class.

As can be seen from this analysis, the two assessment strategies differ on several dimensions. Among the most salient differences are the types of goals they address (knowledge versus application), their language demands (written versus oral), their frequency of occurrence (weekly versus once per semester), and the grading criteria employed (objective versus more subjective). For professors interested in analyzing their own current assessment strategies, a blank analysis worksheet is provided in Appendix A of this chapter.

MODIFYING TRADITIONAL ASSESSMENT PRACTICES

There are a number of relatively minor changes that professors can make in their courses so that their assessment practices are more responsive to the needs of all students, but especially those from culturally diverse groups and other nontraditional students. Four types of changes are discussed in the sections that follow. The first and the last deal with issues of classroom climate and the hidden curriculum: instructor expectations and grading practices. In the middle are suggestions for modifying the two assessment tools most commonly used in college courses, exams, and papers.

Checking Assumptions

When planning a course, professors make assumptions about the skills, knowledge, and attitudes that students will bring with them into the classroom. These assumptions, warranted or not, then translate into expectations for student performance. Dissonance occurs when these expectations are not met. One way to prevent this disjuncture is to gather information early in the term about students' past experiences and current skill levels. There are several ways to go about the task of checking one's assumptions, as the following suggestions illustrate.

1. *Determine whether students have completed prerequisite coursework and other relevant university requirements.* There are many reasons why a student might enroll in a class for which he or she has not met the prerequisites. Some, not convinced of the necessity for prerequisites, treat university regulations quite casually. Others have received poor advising and are not aware of the requirements. Still others select classes that fit into their schedule of work hours, childcare, and/or family commitments.

TABLE 5–1 Analysis of Assessment Strategies

	Task 1	Task 2
Instructional Goal	acquire basic vocabulary in language X	use basic vocabulary in authentic communication situations
Assessment Task	paper-and-pencil quiz	interview native speaker, give oral report on interview
•level	knowledge	application
•input mode	vocabulary lists in each textbook chapter; English words on the quiz	all course activities; class hand-out with suggested interview topics; interview native speaker in language X
•response mode	write language X equivalent for each English word or phrase	give oral report in language X
•aids	none	class notes, text, handouts, dictionary to prepare interview; note cards for report
•language requirements	reading and writing; spelling counts	listening and speaking; in the oral report, students are graded on intelligibility, fluency, word choice, and overall communicative competence
•speed & quantity requirements	15-minute quiz with 20 to 30 items	10-minute interview; 5-minute oral report
•participants	students work alone	students work alone
•location	in class	interview takes place on campus or in the community, oral report in class
Weight	quizzes are 30% of course grade	10% of course grade
Scheduling	every Friday (class meets 3 days a week)	assignment given at midterm; oral reports take place in the last two weeks of the semester
Grading	percentage correct	professor completes rating sheet during report
Feedback	items marked if incorrect; quizzes returned the next class period	rating sheets given to students on the day of the final

Thus, at the start of each class, it is wise to review the required prerequisites with students and explain the rationale behind them and their importance. If deemed necessary, students can be asked to furnish verification of successful completion of prerequisites. As Chapter 4 mentioned, it is also possible to administer a pretest to directly assess mastery of prerequisite skills. Pretests can be structured like traditional course exams with objective and essay questions or, as discussed earlier in this chapter, the instructor may want to experiment with techniques such as concept maps.

2. *Check students' reading and writing skills.* Students' ability to read and understand the class textbook and other required readings can be assessed quite easily with a technique called the cloze procedure (Bormuth, 1968; Jongsma, 1971). This method evaluates silent reading skills. The professor chooses a 250-word passage, leaves the first and last sentences intact, and replaces every fifth word in the rest of the selection with a blank. Students read the passage and write in the missing words (or suitable substitutions). The material is at an appropriate reading level when students are able to predict 50 percent or more of the missing words.

The professor can gain general information about writing skills by asking students to spend 10 or 15 minutes writing a brief essay. An in-class writing sample is preferred because it controls the time variable and the use of aids such as computers with spelling checkers. Having students write a short autobiography or explain their reasons for taking the course also helps the instructor become better acquainted with the students and their experiential backgrounds.

3. *Find out about students' study skills.* Notetaking is one study skill that almost all professors take for granted. As the professor lectures, he or she assumes that students are listening (and comprehending), identifying the major points and important details, and taking careful notes. To assess the accuracy of this assumption, ask students to turn in their class notes at the end of a lecture. Results are always interesting (and often surprising). They also provide almost as much information on the instructor's lecturing style as they do on students' notetaking skills. Angelo and Cross (1993) suggest two other strategies that provide useful information about students' comprehension of lectures and other classroom activities. In the Minute Paper, students write brief answers to two questions at the end of a class session: "What was the most important thing you learned during this class?" and "What important questions remain unanswered?" (p. 148). In a similar activity, Muddiest Point, students describe the portion of the class activity that remains most unclear to them.

Other study skills of interest are strategies for reading and taking notes from the textbook, techniques for preparing for exams, test-taking strategies, amount of time devoted to course activities, and knowledge and use of campus resources such as the library. These skills are best assessed via a student questionnaire. To increase the likelihood of honest answers, questionnaires should be turned in anonymously. With their names on the paper, students might be tempted to exaggerate a bit when asked, "Not counting class time, how many hours do you spend per week on this course?"

4. *Maintain high expectations for all students.* The information gathered about students' skills and knowledge before the course begins is used to modify and improve instruction and to link students with appropriate campus-based resources and support services (Richardson, 1989). Although the professor may need to rethink some of his or her assumptions, this rethinking should not result in lowered expectations. If students lack prerequisite coursework, the solution is enrollment in another class, not lowering standards. If students need to improve their reading skills, they should be advised about options such as peer tutors and study groups, not excused from course reading assignments. As the literature on education and students from diverse groups clearly indicates (e.g., Banks, 1993; Rodriguez, 1991), high expectations play a powerful role in fostering academic success.

Exams

Tests, quizzes, and exams are an accepted part of college coursework. These traditional assessment tools are an efficient method of gathering information about student performance, although they may not be the most effective way, particularly for students for diverse groups. To reduce the barriers that tests present to some students, the measures themselves can be modified and changes can be made in the testing process.

1. *Prepare students for testing.* Test preparation involves informing students about the characteristics of the test they will take and reminding them of (or introducing them to) effective test-taking strategies. Professors can share information about:

- the content that the test will cover (e.g., what chapters or readings are included, a list of the most important terms and concepts, etc.)
- the types and numbers of questions on the test (e.g., 50 multiple-choice items, two essays)
- the weight of each part of the test
- grading procedures (e.g., if partial credit is possible)
- any necessary materials or equipment (e.g., calculator, dictionary, special answer sheet, number 2 pencil)
- strategies for taking the test (e.g., read each item carefully, budget your time, leave difficult items till last, spend the most time on the parts of the test that count the most, and so on)

2. *Create a classroom climate that reduces test anxiety* (Ramsden, 1992). Preparing students for testing is one way to reduce test anxiety; explaining what will occur on test day helps to demystify the assessment process. The general classroom climate is another important contributor. Test anxiety is much less of an issue in classrooms where students feel valued and supported.

3. *Monitor the linguistic complexity of test items.* Test items that are difficult to read pose real problems for some students, particularly those who speak English as a second language. Students are likely to have trouble with sentences that contain double negatives. Long sentences, particularly those with parenthetical remarks (whether those remarks are placed within the confines of parentheses or not), tend to be more difficult to read than shorter sentences. It is also important to point out that abstruse vocabulary and complex sentence structures can have an egregious impact on difficulty level as can sentences which are written in the passive voice. Tests should assess course content, not reading skills. So, avoid double negatives. Write short sentences. Monitor vocabulary level. Use the active voice.

4. *Use performance data to evaluate and revise test items.* After tests are administered and graded, the professor should review and evaluate how students performed on individual test items. When tests are scored by computer, two statistics are typically available for each item: item difficulty and discrimination index. Difficulty level is the percentage of students who pass the item. Brown (1981) recommends a difficulty level of 0.60 to 0.75 (i.e., 60% to 75% of the students respond correctly). When more than half of the students miss an item, it is useful to analyze their responses to determine if a pattern exists. Some professors have a policy of throwing out test items that a majority of the class misses.

The discrimination index compares how well high-scoring students do on an individual test item compared to low-scoring students. The class is divided into two groups, those with total test scores in the upper half and those whose scores fall in the

lower half, and the percentage of students who answer each item correctly is computed separately for each group. The discrimination index is determined by subtracting the percentage of lower-half students from the percentage of upper-half students. The index can be positive or negative. For instance, if one-fourth of the higher-scoring students and two-thirds of the lower-scoring students are successful on a particular item, its discrimination index is negative ($0.25 - 0.67 = -0.42$). Gay (1985) suggests that items show adequate discrimination when the index is +0.30 or greater.

Similar analysis techniques can be used when tests are scored by hand. At minimum, professors should look for questions that posed significant challenges for large numbers of students. High error rates can be caused by unclear directions, ambiguous wording, or even an incorrect answer key.

5. *Consider using a variety of testing options.* This strategy requires the professor to re-think his or her current practices and consider alternative methods of testing. Possibilities include:

- *Give power tests instead of speed tests* (Boone, Kaiser, & Litowitz, 1988). Is speed an important component of the knowledge or skills the test is designed to measure? If not, reduce the number of test items, increase the time allowed for the test, and/or consider giving take-home rather than in-class tests.
- *De-emphasize rote memory.* Is rote memory the cognitive skill under assessment, or is it a higher level skill such as problem-solving? If memory is not the major concern, consider options such as open-book and take-home exams.
- *Vary the types of test questions.* There are a number of choices for objective-type questions: true-false, multiple-choice, matching, completion (fill-in-the-blank), short answer, and problems (McLoughlin & Lewis, 1994). Essay questions are another option. Ramsden (1992) advocates less use of objective items and more emphasis on essays and other types of questions that require students to demonstrate their understanding of the course material.
- *Allow the use of aids.* If spelling skills are not the target of assessment, let students use dictionaries. Likewise, if computational skills are not being evaluated, students should be permitted to use calculators.
- *Replace two or three big tests with several smaller ones.* Students are under less pressure when assessment occurs frequently. They also have more opportunity to benefit from the feedback from one test to prepare for the next.
- *Provide ways for students to recover from one bad grade.* Among the possibilities are "forgiving" the lowest grade, giving make-up tests, and accepting additional work for extra credit.

6. *Be prepared to work with students with disabilities on accommodations for testing.* Students with learning disabilities often request extra time for exams (most typically time and a half). Students who are blind will need a braille copy of the test or a reader to read the test aloud to them. Other common testing adaptations are use of a dictionary, electronic spelling aid, or word processor; a scribe to record the student's oral responses on the test form or answer sheet; and test administration in a quiet environment free from distractions.

Papers

Like exams, term papers and other types of written assignments are a common feature of university coursework. Papers draw upon a different set of student skills from those required by exams, and so the strate-

gies used to adapt them are somewhat different. However, the first consideration in either case is preparation of the student.

1. *Communicate expectations clearly.* The professor should explain the task and the procedures for accomplishing it. If the assignment is a research paper, for instance, it is helpful to review the various steps in the process (selecting a topic, searching the literature, reading and notetaking, etc.) and to suggest timelines for each step. Length, style, and formatting requirements should be discussed as well as the criteria for grading the final product. If a rating sheet, checklist, or feedback form will be used in grading, give students a copy when the paper is assigned. Students are more likely to spend their time wisely if they know that the summary and conclusions section of the paper is worth 25 points whereas the introduction and problem section is worth only 7. Also helpful are model papers from previous semesters.

2. *Encourage the use of aids and resources.* Students should be encouraged to use dictionaries, thesauruses, word processors with spelling and grammar checkers, and other pertinent aids as they write their paper. In addition, they should be made aware of campus resources where they can obtain these aids and other types of assistance. Students may not be familiar with campus services such as a word processing lab, computer literature searches, reference librarians, interlibrary loan, or term-paper clinics run by the academic skills center.

3. *Consider making peer editing part of the assignment* (Harrison, 1992). Peer editors read and critique each other's work. Typically, the professor structures the editing process by providing students with a list of areas to consider or questions to answer in their review. For instance, peer editors should provide feedback not only on the weaknesses of a paper but also its strengths. Peer reviews are usually written rather than oral and, in some classes, they are graded as part of the assignment.

4. *Break large writing tasks into smaller chunks.* Big assignments such as term papers can be broken down into a series of smaller tasks. For example, students could turn in a one-paragraph description of their topic in week 3 of the semester, their note cards in week 6, an outline of the paper in week 9, and a rough draft in week 12. These intermediate tasks may be graded or simply critiqued. In either case, students benefit because they receive feedback on their work-in-progress.

5. *In grading, separate form and content issues.* Instructors are influenced by the appearance of a paper (Good & Brophy, 1987) and by any errors it contains in spelling, grammar, punctuation, and the like. One way to counteract this bias is to grade the content of a paper and its form separately. For instance, if content is the major concern, it can be weighted so that it accounts for 90 percent of the grade and form is the remaining 10 percent.

Grading Practices

Although grading has been mentioned in relation to modifications for exams and papers, there are some general considerations that apply to any type of assessment task. Each relates to classroom climate and the messages professors send to students as they evaluate their work.

1. *Be aware of the impact of your written and verbal comments.* Imagine spending over a month writing a paper, then getting it back bloodied with red ink and without a single positive comment. This experience is all too common in college classrooms. It takes only a few seconds more to add a few words of encouragement when grading a paper. And it is just as easy to grade a test

with green or purple ink, colors less associated with failure. Similar caveats apply to verbal feedback. Words, whether spoken or written, have the power to wound or to spur students on to greater achievements.

2. *Do not encourage competition among students.* Competition is not equally valued by all cultures. As Chapter 4 pointed out, some cultures and many women prefer cooperative interactions to competitive ones (Lynch & Hanson, 1992). Professors can inject a competitive spirit into grading in big ways and in small ones. Grading on a curve is a major endorsement of competition. Contrasting the performance of different groups (the men in the class versus the women, the English majors versus the Education majors) when discussing results of a test or assignment is more subtle, yet equally distasteful to some groups.

3. *Be cautious about singling out students for individual recognition.* No one likes to be singled out for criticism. In some cultures, individuals are equally uncomfortable being singled out for praise.[1] Thus, it is best to confine laudatory remarks to written comments or private conversations, at least until students' preferences on this issue can be determined. One simple way of discovering students' wishes is to distribute a brief questionnaire at the start of the semester or quarter. That questionnaire might ask students to comment on their preferences for feedback and recognition and to share their views on other topics such as the types of instructional activities to be used in the class and the scheduling of office hours for the instructor and any teaching assistants.

ALTERNATIVE ASSESSMENT PRACTICES

No course can be considered multiculturally infused if it utilizes only traditional assessment procedures. As professors move their courses along the course-change continuum described in Chapter 2—from exclusive to inclusive to transformed—they must begin to develop new ways of thinking about the assessment process. This final section of the chapter provides an introduction to some of those new perspectives.

Support Student Learning with Peer, Campus, and Community Resources

One professor, by himself or herself, cannot hope to meet the needs of every member of every class. This caution does not mean, however, that those needs cannot be addressed. To successfully accommodate students' needs, two things must occur: an expansion of the concept of "classroom" to include places outside the classroom walls, and a broadening of the concept of "teacher" to include persons other than the professor.

There are many campus and community resources that can assist nontraditional students to succeed in higher education. Several have already been mentioned earlier in this chapter and in previous chapters. Richardson (1989), in a report of successful strategies for increasing the scholastic success of students of color, identifies a range of effective campus-based services related to the academic program:

- Tutoring is widely available to students who need it.
- Assistance with reading, writing and math skills is available on a walk-in basis.
- Instruction in study skills, note-taking and test preparation is provided to all students as needed.
- Students who follow nontraditional patterns of attendance have access to an educational service center that provides counseling, developmental coursework, tutoring, critical reading and library research skills, time management and study skills. (p. 12)

Another source of assistance—one that is often overlooked—is the student members of the class. Students can help each other in a variety of ways: as tutors, academic advisors, mentors, editors, consultants, and sources of encouragement and support. Rodriguez (1991) describes a program at Berkeley where African American students were brought together in study groups. The program was started because Uri Treisman, its founder, noticed that "Black students [who were failing] studied alone, even if for many long hours, while Asian students, who were passing, studied in groups. Through study groups Asian students tutored, criticized and supported each other" (p. 7). (Chapter 8 describes an adaptation of Treisman's model for Latino, Black, and other underrepresented students in freshman calculus.)

Cooperative learning is another strategy that involves peers as teachers. Students work collaboratively in groups on common tasks or to solve common problems (see Chapter 7 for a detailed discussion). Cooperative learning can extend into assessment. A colleague of mine teaches a doctoral level statistics course in which students work in cooperative groups throughout the semester. On the nights of the midterm and final, the students meet as groups to discuss the test questions before each writes his or her individual answers (Santa Cruz, 1993).

Contextualize Learning and Its Assessment

In many cultures, knowledge is considered worth learning only when it is perceived as relevant to real-world issues and the solution of real-world problems. In addition, in many cultures, children are reared to value interactions with people above interactions with things.[2] These value systems carry with them three important implications. First, learning (and its assessment) should be contextualized so that students perceive it as both relevant and important. When studying learning in a basic psychology course, for example, students might be asked to provide examples of how they have applied course content in their own studying (Angelo & Cross, 1993). Second, whenever possible, learning and its assessment should include opportunities for social interaction. Cooperative learning strategies provide these opportunities as do course assignments in which students interact with others outside the classroom walls. For example, students in a political science course might participate in a local election; students in a history course might interview community members about their recollections of and participation in the civil rights movement. Third, learning and its assessment are tied to real-world issues, and students are encouraged to use critical thinking skills to apply course content to the solution of real-world problems. For instance, as a culminating activity for a course on child development, students might be asked to draft a set of recommendations to improve the well-being of children in their community, state, or the nation as a whole.

Use Multiple Evaluation Techniques

Contemporary theories of intelligence recognize that ability is not a unidimensional trait. Gardner (1987), for example, describes seven types of intelligence: linguistic, logical-mathematical, spatial, musical, bodily kinesthetic, interpersonal, and intrapersonal. In Sternberg's triarchic theory (Sternberg, 1984; Sternberg, Okagaki, & Jackson, 1990), intelligence is contextualized. It includes not only an individual's mental mechanisms for processing information but also the tasks and situations in which that processing occurs and the sociocultural context in which the individual lives.

Theories such as these call attention to the multidimensional nature of human ability. Strengthening this position is the information available on the ways in which some women, individuals with disabilities, and persons of color prefer to engage in cognitive tasks. For example, as earlier chapters have described, some groups prefer holistic approaches to problem solving rather than analytic approaches.[3]

Student achievement can be enhanced by providing multiple ways for students to demonstrate their mastery of course competencies. There are many ways to vary the assessment process, and the model described in the first part of this chapter outlines a number of dimensions to consider. For example, the most typical input modes for assessment tasks, reading the textbook and listening to lectures, are routinely altered for students with disabilities (Lewis, 1993). Those who cannot read because of a vision loss or a learning disability listen to Talking Books or books on tape. Those who cannot hear watch an interpreter as he or she translates the lecture into sign language. Similar types of alternatives can be made available to all students. Among the possibilities are reading related fiction as a supplement to the textbook, watching a movie or video (with captions, if the class includes students with hearing losses or those for whom English is a second language), conducting an interview or observation, and participating in a community service activity.

The response modes for assessment tasks can also be altered. For example, a traditional term paper assignment might be replaced with a reading log or a series of abstracts annotated with students' personal reactions. A traditional exam might be replaced with a concept map exercise where students must fit several key concepts into a map prepared by the instructor (Seymour, 1993). In addition, writing can be replaced by other response modalities such as speaking, artistic expression, music, or performance art. Many years ago, when I was struggling to teach undergraduate students about learning disabilities, I asked them to develop a personal definition of this condition using any medium they wished. The "papers" I remember best are a poem, an original piano composition, and a macramé wall hanging.

Performance-based assessments allow students to demonstrate in some concrete way their mastery of course material. For example, in mathematics, a student might be asked to solve a complex problem and explain the process used to arrive at the solution. In education, a student might videotape a lesson he or she has taught to document competence in instructional techniques. Banta (1992) provides several other examples including an in-basket activity where students in dietetics must respond to mail received by a hospital dietitian and an activity for theater majors which requires them to view a videotape of a portion of a play and, depending on their specialization, critique the acting, directing, or set design.

Portfolio assessment is a technique that can incorporate a range of response modes (Abruscato, 1993; Feuer & Fulton, 1993; Grady, 1992; Wolf, 1991; Worthen, 1993). Portfolios contain samples of student work collected over time (e.g., over the duration of a course or the completion of an academic major) to document progress and record accomplishments. As Banta (1992) comments, portfolios can contain a variety of student products:

> ...course assignments, research papers, materials from group projects, artistic productions, self-reflective essays, correspondence, and taped presentations. Student performances can be recorded using audio- or videotapes. Potential materials for the cassette-recorded

> portfolio are speeches, musical performances, visual arts productions, foreign language pronunciation, group interaction skills, and demonstrations of laboratory techniques or psychomotor skills. (p. 1688)

Two characteristics make portfolio assessment a good choice for students from diverse groups: students themselves are responsible for selecting the work samples that go into the portfolio (with some initial guidance from the professor), and one of those samples is typically a self-evaluation. When logistically feasible, portfolios should be reviewed as part of a professor-student conference.

Empower Students to Make Choices

It is not enough to incorporate many types of assessment tasks into a course; students also must be able to choose among them. What we as professors see as benign assignments may be perceived as insulting, degrading, or extremely distasteful by some students. A colleague of mine tells a story from her high-school years about being assigned to compete against her best friend in a debate. As Japanese American young women, they were horrified with being forced into competition against one other.

Choice-making is a means of empowerment. Giving students options acknowledges their competence and the value of their preferences. It also helps them begin to assume responsibility for setting their own goals, directing their own learning, and participating in the evaluation of that learning. That beginning is an important part of the true commencement: the start of lifelong learning.

Take out a blank sheet of paper and draw a concept map of "classroom assessment." Compare this map with your original version to evaluate how your thinking has changed. If you would like further information about current trends and issues in classroom assessment, see the work of Banta (1992) and Ewell (1991). For specific suggestions and examples of classroom assessment techniques, consult the handbook prepared by Angelo and Cross (1993).

ENDNOTES

1. Authors such as Cheng (1987), Joe and Malach (1992), Kitano (1973), Pepper (1976), and Walker (1988) report that persons with Asian roots and persons with Native American roots tend to value cooperation rather than competition and achievement of the group rather than achievement of the individual. However, it is extemely important to recognize that descriptions of cultures are by their nature generalizations. Lynch (1992) warns: Culture is only one of the characteristics that determine individuals' and families' attitudes, values, beliefs, and ways of behaving...Assuming that culture-specific information...applies to all individuals from the cultural group is not only inaccurate but also dangerous—it can lead to stereotyping that diminishes rather than enhances cross-cultural competence. When applying culture-specific information to an individual or family, it is wise to proceed with caution. (p. 44)
2. For example, authors such as Franklin (1992), Willis (1992), and Zuniga (1992) suggest that persons with African American roots and persons with Latino roots place high value on the importance of the family, interpersonal interactions, and the solution of real-world problems, particularly those related to social issues such as inequality.
3. Chan (1992), for example, notes that persons identifying with the traditional Asian culture favor "contemplative, circular thinking" whereas those identifying with the dominant Anglo-European culture favor "analytic linear thinking" (p. 251).

TABLE CHAPTER 5, APPENDIX A Worksheet for Analyzing Assessment Strategies

	Task 1	Task 2	Task 3
Instructional Goal			
Assessment Task			
• level			
• input mode			
• response mode			
• aids			
• language requirements			
• speed & quantity requirements			
• participants			
• location			
Weight			
Scheduling			
Grading			
Feedback			