
PART ONE

Understanding Assessment

COPYRIGHTED MATERIAL

What Is Assessment?

Some Valuable Ideas You'll Find in This Chapter

- Students learn best—and assessment works best—when education is a purposeful, integrated, collaborative experience.
- Teaching to the test may not be such a bad thing.
- Assessment is research, but it is not traditional empirical research.
- Grades alone may not tell us much about student learning, but the grading process can yield a wealth of valuable information.

Oral and written examinations have been part of education for hundreds of years, but only in the past century have the theory and science of assessment been studied systematically. Because the assessment of student learning in higher education is relatively new compared to many other fields of study, and because it has been undertaken by people from disciplines with widely differing orientations, the vocabulary of assessment is not yet standardized. (This chapter, for example, discusses several ways that the term *evaluation* is used.) This book therefore begins by defining assessment and distinguishing it from some related concepts.

Many assessment practitioners, notably Thomas Angelo (1995), have put forth definitions of assessment. Table 1.1 summarizes their work.

The four steps in Table 1.1 do not represent a once-and-done process but a continuous four-step cycle (Figure 1.1). In the fourth step, assessment results are used to review and possibly revise approaches to the other three steps, and the cycle begins anew.

Table 1.1. What Is Assessment

Assessment is the ongoing process of:

Establishing clear, measurable expected outcomes of student learning

Ensuring that students have sufficient opportunities to achieve those outcomes

Systematically gathering, analyzing, and interpreting evidence to determine how well student learning matches our expectations

Using the resulting information to understand and improve student learning

While the term *assessment* can be used broadly—we can assess any goal or outcome in any discipline or any activity—in this book, the term refers to the assessment of student learning.

What Is the Difference Between Traditional and Current Approaches to Assessment?

How are today's approaches to assessment different from the oral and written examinations that faculty have been conducting for centuries? Table 1.2 summarizes some key differences.

An important difference between contemporary and traditional thinking about assessment is that under contemporary approaches, assessment is viewed as part of an integrated, collaborative learning experience. Students learn better when their college experiences are not collections of isolated courses and activities but are purposefully designed as coherent, integrated learning experiences in which courses and out-of-class experiences build on and reinforce one another (see Table 18.3). Indeed, Gerald Graff (2008) has noted that successful colleges stress collaboration over “individual teaching brilliance” and that students find unrelated courses

Figure 1.1. Teaching, Learning, and Assessment as a Continuous Four-Step Cycle

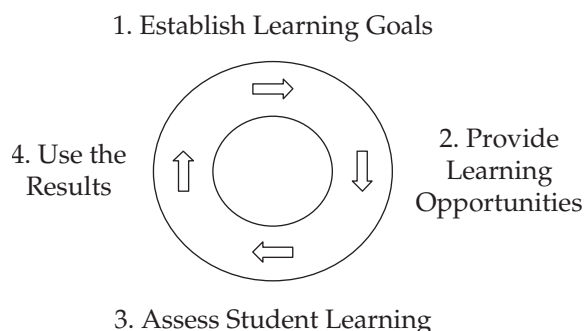


Table 1.2. Contemporary Versus Traditional Ways of Thinking About Assessment

<i>Contemporary Approaches: Assessment is . . .</i>	<i>Traditional Approaches: Assessment is . . .</i>
Carefully aligned with goals: the most important things we want students to learn (Chapter Eight)	Planned and implemented without consideration of learning goals, if any even exist
Focused on thinking and performance skills (Chapter Eight)	Often focused on memorized knowledge
Developed from research and best practices on teaching and assessment methodologies (Chapter Eighteen)	Often poor quality because faculty and staff have had few formal opportunities to learn how to design and use effective assessment strategies and tools
Used to improve teaching and learning as well as to evaluate and assign grades to individual students (Chapters Four and Eighteen)	Used only to evaluate and grade individual students, with decisions about changes to curricula and pedagogies often based on hunch and anecdote rather than solid evidence
Used to tell our story: what makes our college or program distinctive and how successful we are in meeting students' and societal needs (Chapter Seventeen)	Not used to tell that story; stories are told through anecdotes about star students rather than broader evidence from representative students

confusing. When students can see connections among their learning experiences, their learning is deeper and more lasting.

The value of education as an integrated, collaborative experience has several important implications for teaching and assessment:

- *Integrated learning goals.* There should be appropriate relationships among institutional, program, and course learning goals. This is discussed in Chapter Eight.
- *Curricular alignment.* Curricula should be designed to ensure that every student, regardless of the particular choices he or she makes in choosing a course of study, has ample opportunity to achieve every key institutional and program learning goal. This is discussed in Chapters Three and Seven.
- *Collaboration.* Learning goals, curricula, and assessments should be designed through collaboration across the college community. This is discussed in Chapter Five.
- *Embedded assessments.* An important side benefit of providing integrated learning experiences is that student learning assessments can be similarly integrated. Assessments that are embedded into individual courses (Chapter Two) can often provide information on student achievement of program goals, general education goals, and institutional goals.

What Are the Differences Among Course, Program, General Education, and Institutional Assessments?_____

Student learning takes place in many venues:

- Individual courses
- Academic programs, including undergraduate and graduate degree programs, certificate and other nondegree programs, and noncredit programs
- General education core curricula
- Cocurricular programs and student life programs designed to promote student learning and development
- Cohort-based programs and other special programs designed to enhance student learning, such as:
 - First-year experiences
 - Learning communities
 - Service-learning programs
 - Developmental education programs
 - Tutoring programs
 - Honors programs
 - Programs for at-risk student cohorts
 - Study-abroad programs

Assessment at the Course Level

Assessment in individual courses is typically based on the tests and assignments that contribute to the grading process. Under the contemporary approaches to assessment listed in Table 1.2, assessment at the course level means not just assigning individual grades but also reflecting on how well students as a whole are achieving the course's key learning goals. This is done by aggregating assessment results (Chapter Sixteen), such as by counting how many students answered each test question correctly. This takes time, of course, but can often be accomplished by looking at the results of just a few key assignments, generally those completed toward the end of the course.

Course assessment becomes more complicated if several faculty members are teaching multiple sections of the same course and using different tests, assignments, and other grading criteria. If the course is a prerequisite to further study or to a career, all sections should share a core of common course learning goals (Chapter Eight) that are essential to future success. (Individual faculty members may, of course, address additional goals of their

own choosing beyond this common core.) It can be very helpful to come up with some common test questions (Chapter Eleven) or assignments (Chapter Ten) and develop common criteria (Chapter Nine) to evaluate these common goals.

Assessment at the Academic Program Level

Because an academic program should be integrated and greater than the sum of its parts—that is, more than a collection of courses—it may have goals and assessments that are broader than those of its courses. A course assessment might examine whether students can solve a specific kind of problem, for example, while a program assessment might examine whether students can design appropriate approaches to solving a variety of problems in the discipline.

Assessment at the academic program level can take place in a variety of ways.

Embedded course assignments. Course assignments, especially those completed toward the end of a program, can be assessed for achievement of key program goals as well as course goals. A paper that a student writes in an advanced course, for example, can show not only what the student has learned in that course but also the writing skill that she has developed throughout the entire program of study. Embedded assessments are discussed further in Chapter Two.

Capstone experiences. These are holistic activities that students are required or encouraged to complete as they approach the end of their program. They include theses, dissertations, oral defenses, exhibitions, performances, presentations, and research projects. Capstones help students synthesize their learning by tying together the various elements of their program and seeing the big picture. Capstones thereby promote deep, lasting learning (Table 18.3). These experiences provide a wonderful venue for program assessment because they provide a holistic portrait of what students have learned throughout their program. A senior project might be evaluated—perhaps through a rubric—for such program goals as written communication skills, critical thinking skills, information literacy skills, and research skills. If students make presentations on their projects, the presentations can be evaluated for oral communication skills as well.

Field experiences. Many programs require students to participate in an internship, practicum, service-learning activity, student teaching assignment, or some other capstone experience in the field. If

these experiences give students opportunities to practice applying the knowledge and skills they've learned in the program to real-life situations, their supervisors' ratings of their performance can be powerful evidence of the overall success of a program in achieving its major learning goals. Rubrics make it easy for supervisors to provide this information (see Exhibit 9.5).

Portfolios. Academic programs increasingly require students to compile a portfolio that demonstrates what they have learned throughout the program. Portfolios can draw a rich, full picture of student learning, but they can also be complicated and time-consuming to implement. Portfolios are discussed in Chapter Thirteen.

Published tests. Some programs require or encourage students to take a published test of what they have learned. Published tests are discussed in Chapter Fourteen.

Assessment in General Education Core Curricula

As with academic programs, a good general education core curriculum is greater than the sum of its parts. It has overarching goals and because those goals are integrated, they are the hallmark of every undergraduate's education at the college. Those goals are addressed repeatedly in multiple rather than single general education courses, and the assessment of general education focuses on those goals.

Catherine Palomba and Trudy Banta (1999) have offered three general approaches to assessing general education learning goals. One approach is to let faculty identify their own embedded assessments (Chapter Two) of the general education courses they teach. This approach gives faculty the greatest sense of ownership and may therefore generate the most useful results. But this approach makes it difficult to aggregate the results and get an overall picture of how well students are achieving general education goals across the entire college.

Another approach is to use a collegewide assessment, perhaps a portfolio, a published test of general education skills, or a capstone requirement of the general education curriculum. While this makes it easy to get an overall picture of student learning, faculty may not be able to see how the results relate to their classes or how to use the results to improve student learning. Another challenge is that college-wide assessments are often add-on assessments, and student motivation may be an issue (Chapter Two).

The third approach is for faculty teaching courses in a group of related disciplines or subjects to identify a common assessment strategy. Faculty teaching science laboratory courses, for example, might agree to include certain key elements in their rubrics for students' lab exercises.

Unless a consistent approach is mandated by an accreditor, state agency, or the like, there is usually no need to use just one approach throughout the general education curriculum. Writing might be assessed through a collegewide portfolio requirement, fine arts faculty might each develop their own assessments of the general education goal for creativity, and the science faculty might agree to include a common set of test questions on their final exams.

Assessment in Cocurricular, Student Life, and Cohort-Based Programs

Student learning takes place outside as well as within the curriculum: in first-year experiences, learning communities, other cocurricular programs, and student life programs. Wherever student learning and development are supposed to happen, there should be goals for that learning and assessments to see how well students are achieving those goals.

Many of these programs have goals to develop attitudes, values, and the like. Strategies to assess these kinds of goals are discussed in Chapter Twelve.

Assessment at the College or University Level

Many colleges have overarching learning goals for all students, regardless of major. Frequently these institutional learning goals are delivered through the general education core curriculum. If this is the case, assessment at the institutional level is synonymous with general education assessment.

But some college mission statements and strategic goals articulate institutional learning goals that are not addressed systematically in the general education curriculum. A faith-based college, for example, may have goals related to spirituality that students develop through participation in extracurricular activities rather than—or in addition to—the general education curriculum. Some colleges have a distinctive goal that students develop through both the general education curriculum and their academic major field of study. Hamilton College in Clinton, New York, for example, characterizes itself as “a national leader for teaching students to write effectively” (Trustees of Hamilton College, 2008), and it emphasizes this skill throughout its curricula.

In these circumstances, institutional learning assessment goes beyond general education assessment, but the principles for assessing general education curricula apply.

Responsibility for institutional learning assessment is often shared—or should be shared—not only among faculty but also with student development staff. Interpersonal skills are an example of an institutional goal that might be developed and assessed in the general education curriculum, majors, and student development programs. This requires communication and collaboration, as discussed in Chapter Five.

What Is the Difference Between Assessment and Grading?_____

Obviously there is a great deal of overlap between the tasks of grading and assessment, as both aim to identify what students have learned. A key difference is that grades focus on individual students, while assessment focuses on entire cohorts of students and how effectively everyone, not an individual faculty member, is helping them learn. Grades alone are usually insufficient evidence of student learning for assessment purposes (Johnstone, Ewell, & Paulson, 2001) for several reasons:

Grades alone do not usually provide meaningful information on exactly what students have and have not learned. We can conclude from a grade of B in an organic chemistry course, for example, that the student has probably learned a good deal about organic chemistry. But that grade alone cannot pinpoint what aspects of organic chemistry she has and has not mastered.

Grading and assessment criteria may (appropriately) differ. Some faculty base grades not only on evidence of what students have learned, such as tests, papers, presentations, and projects, but also on student behaviors that may or may not be related to course learning goals. Some faculty, for example, count class attendance toward a final course grade, even though students with poor attendance might nonetheless master course learning goals. Others count class participation toward the final grade, even though oral communication skills aren't a course learning goal. Some faculty downgrade assignments that are turned in late. These practices can all be appropriate classroom management strategies and grading practices, but they illustrate how grades and assessment standards might not match. A student who does not achieve major learning goals might nonetheless earn a fairly high grade by playing by the rules and fulfilling other less important grading criteria.

Conversely, a student who achieves a course's major learning goals might nonetheless earn a poor grade if she fails to do the other things expected of her.

Grading standards may be vague or inconsistent. While many faculty base assignment and course grades on carefully conceived standards, grades can be inadequate, imprecise, and idiosyncratic, as Thomas Angelo has pointed out in the Foreword to the first edition of *Effective Grading* (Walvoord & Anderson, 1998). Faculty may say they want students to learn how to think critically but then base grades largely on tests emphasizing factual recall. Faculty teaching sections of the same course may not agree on common standards and might conceivably award different grades to the same student performance on the same assignment. Sometimes individual grading standards are so vague that a faculty member might, in theory, award an A to an essay one day and a B to the identical essay a week later.

Grades do not reflect all learning experiences. As *Greater Expectations* (Association of American Colleges and Universities, 2002) points out, grades provide information on student performance in individual courses or course assignments. They do not provide information on how well students have learned key competencies, such as critical thinking or writing skills, holistically over an entire program. Grades also do not address what students have learned from ungraded cocurricular activities.

Do grades have a place in an assessment program? Of course they do! Grades can be useful, albeit indirect, evidence (Chapter Two) of student learning. They can be useful *if* the grades are based on direct evidence of student learning (Chapter Two) such as tests, projects, papers, and assignments that are clearly linked to major learning goals through test blueprints (Chapter Eleven) or rubrics (Chapter Nine). *Effective Grading* (Walvoord & Anderson, 1998) gives a plethora of practical suggestions on how to tie grades more closely to explicit learning goals.

What Is the Difference Between Assessment and Teaching to the Test?

In a way, good assessment *is* teaching to the test. Assessment is part of a process that identifies what we want students to learn, provides them with good opportunities to learn those things, and then assesses whether they have learned those things. In

other words, good assessment assesses “what matters most” (Angelo, 1999, p. 3).

Teaching to the test gets a bad name when tests measure something other than what we value, either because someone else has told us what to assess or because our own tests measure relatively trivial learning. As Lee Shulman (2007) has pointed out, “Assessments must be designed so that the tests are *worth* teaching to” (p. 24).

What Is the Difference Between Assessment and Evaluation?_____

Evaluation is defined in a variety of ways. One perspective equates it with judgment: evaluation is using assessment information to make an informed judgment on such things as:

- Whether students have achieved the learning goals established for them
- The relative strengths and weaknesses of teaching and learning strategies
- What changes in goals and teaching-learning strategies might be appropriate

Evaluation defined this way is the last two steps of the assessment process described at the beginning of this chapter: interpreting assessment evidence (part of step 3) and using the results (step 4). This definition points out that assessment results alone only guide us; they do not dictate decisions to us. We use our best professional judgment to make appropriate decisions. This definition of evaluation thus reinforces the ownership that faculty and staff have over the assessment process.

A second conception of evaluation is that it determines the match between intended outcomes (step 1 of the assessment process) and actual outcomes (step 3 of the assessment process). Under this definition, the assessment of student learning and the evaluation of student learning could be considered virtually synonymous.

A third conception of evaluation is that it investigates and judges the quality or worth of a program, project, or other entity rather than student learning. We might evaluate an anthropology program, employee safety program, alumni program, or civic project designed to reduce criminal recidivism. Under this definition, evaluation is a broader concept than assessment. While assessment focuses on how well student learning goals are achieved, evaluation addresses how well all the major goals of a program are achieved. An anthropology program, for example, might have goals not only for student learning but also to conduct anthropological

research, provide anthropological services to local museums, and conduct its affairs in a cost-effective manner. An evaluation of the program would consider not only student learning but also research activities, community service, and cost-effectiveness. Program reviews, discussed below, are an increasingly popular tool for evaluating academic programs.

What Is the Difference Between Assessment and Faculty Evaluation?_____

While faculty evaluations examine the impact and effectiveness of individual faculty members, assessments of student learning evaluate the collective impact of faculty, staff, and the resources that support them on students in an entire program or an entire college or university.

As discussed in Chapter Eighteen, one of the fastest ways to kill an assessment effort is to use the results to evaluate individual faculty, especially if disappointing results are used to penalize faculty. But end-of-course evaluation forms completed by students might provide useful assessment information if they are aggregated across faculty. A question about opportunities for interaction with other students may provide useful, albeit indirect (Chapter Two), program assessment information, for example.

What Is the Difference Between Assessment and Research?_____

Lee Upcraft and John Schuh (2002) have noted that assessment differs from traditional research in its purpose and therefore in its nature. Traditional empirical research is conducted to test theories, while assessment is a form of action research, a distinct type of research whose purpose is to inform and improve one's own practice rather than make broad generalizations. Peter Ewell (2002) has called this a craft-based rather than scientific approach. The four-step assessment cycle of establishing learning goals, providing learning opportunities, assessing student learning, and using the results mirrors the four steps of action research: plan, act, observe, and reflect.

Assessment, like any other form of action research, is disciplined and systematic and uses many of the methodologies of traditional research. But most faculty and staff lack the time and resources to design and conduct rigorous, replicable empirical research studies with impartial distance. They instead aim to keep the benefits of assessment in proportion to the time and resources

devoted to them (Chapter Six). If you take the time and effort to design assessments reasonably carefully and collect corroborating evidence, your assessment results may be imperfect but will nonetheless give you information that you will be able to use with confidence to make decisions about teaching and learning. Chapters Three and Sixteen discuss strategies for doing this.

What Is the Difference Between Assessment and Program Review?

Program review is a comprehensive evaluation of an academic program that is designed both to foster improvement and demonstrate accountability (Suskie, 2006). Program reviews typically include a self-study conducted by the program's faculty and staff, a visit by one or more external reviewers, and recommendations for improvement based on the conclusions of the self-study and the reviewer. Program reviews are sometimes conducted to meet the requirements of an accreditor or state system and sometimes simply because institutional leaders find the concept appealing.

Program reviews can be a useful tool for improvement or a meaningless paper-pushing exercise, depending on how they are designed. Robert Shirley and J. Fredricks Volkwein (1978) have suggested that program reviews focus on quality, along with need, cost, and cost-effectiveness. But what is quality? As Chapter Four discusses, colleges and universities have traditionally defined quality based on inputs, including resources (such as faculty credentials and library holdings), students (such as their high school preparation), and activities (such as faculty scholarship and community service).

But quality and effectiveness increasingly are defined by how well a program is achieving its goals. Under this view of quality, a program review should focus on collecting and examining evidence of goal achievement. Because student learning is a fundamental goal of any academic program, student learning assessment should be a primary component of the program review process.

The value of program reviews depends not only on their focus but also on how they are used. At some colleges, completed program reviews are submitted, filed, and forgotten, making them a pointless exercise. But other colleges use program reviews—including assessment results—to develop plans for advancing programs and allocating resources to support achievement of those plans. Chapter Eighteen discusses using assessment results to inform planning and budgeting decisions.

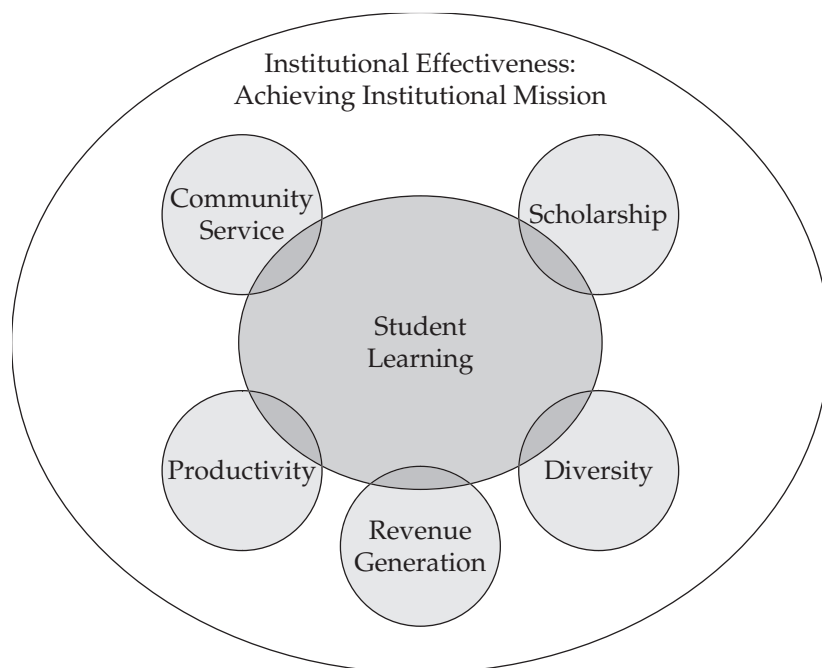
What Is the Difference Between Assessing Student Learning and Assessing Institutional Effectiveness?

Accreditation organizations, governing boards, legislators, and other audiences increasingly ask colleges and universities to assess institutional effectiveness as well as student learning. What's the difference? Institutional effectiveness is how well a college or university is achieving its mission and major strategic goals. Since student learning is the heart of most college missions, the assessment of student learning is a major component of the assessment of institutional effectiveness (Figure 1.2).

But institutional effectiveness goes further, addressing other aspects of college mission, perhaps research and scholarship, community service, building a diverse community, or modeling certain values. Institutional effectiveness also addresses progress in achieving major institutional goals, which might include offering financial support to those without sufficient means to attend college, providing facilities and infrastructure that promote student learning, or developing collaborative partnerships with basic education.

Assessing institutional effectiveness thus involves assessing not only student learning but also each of these other collegewide aims through the same four-step cycle. Assessing a college mission

Figure 1.2. The Relationship of Student Learning to Institutional Effectiveness



of community service begins, for example, by developing a clear statement of the major goals of the college's community service efforts (step 1). Programs designed to achieve those goals are then implemented (step 2). The programs are next assessed to see whether they are achieving their major goals (step 3). For example, a goal to provide cultural programming to the local community might be assessed by counting attendance at cultural events and perhaps surveying local residents on how well the college meets their cultural interests. Finally, the results are used to modify the college's community service goals, programs, or assessment strategies (step 4), and the cycle begins anew.

What Is the Difference Between Assessment and Performance Indicators?

Performance indicators are quantitative measures (Chapter Two) summarizing assessment results for student learning or other aspects of college performance that are distilled down to single numbers. Student retention and graduation rates, job placement rates, racial/ethnic enrollment breakdowns, financial ratios, and student-faculty ratios are all examples of performance indicators. Busy college leaders, board members, and government policymakers often want performance indicators because they are quickly digested, even though they may present an incomplete, if not distorted, picture of effectiveness and quality.

Most performance indicators used today are not measures of student learning because most student learning assessment results are complex and cannot be distilled down to a single number. Rubrics (Chapter Nine) and surveys (Chapter Twelve), for example, usually yield results on multiple criteria.

But some student learning assessment results could be viewed as performance indicators. The percentage of students who earn at least a minimally acceptable overall score on a rubric or test, who give a particular answer to a single survey question, or who earn a certain score on a single rubric criterion are examples of possible performance indicators.

Performance indicators greatly increase the temptation to make quick judgments and decisions based on just one assessment result. As discussed in Chapter Three, assessment results are always imprecise, and no decision should be based on the results of a single assessment.

But we cannot ignore increasingly vocal calls for simple, clear assessment results that public audiences such as employers and policymakers can easily absorb and understand. Until the

education community develops effective, compelling ways to summarize and communicate student learning assessment results, the temptation to look on some results as performance indicators will remain. Chapters Four and Seventeen discuss the challenges of sharing assessment results with public audiences.

What Is the Difference Between Assessment and Accountability?_____

Assessment is the act of evaluating student learning; accountability is using the results of assessment to demonstrate the quality of a program or college to concerned audiences. Chapter Four discusses the twin purposes of assessment: using results to improve teaching and learning and using results to be accountable to internal and external audiences.

Time to Think, Discuss, and Practice _____

1. Think of an assignment in a course you have taught or taken.
 - Did the assignment help students learn important goals or relatively unimportant goals of the course?
 - Might the completed assignment be evidence of student achievement of a program or college learning goal? Which goal?
2. Are the academic programs at your college required to undergo any periodic program review?
 - What are the guidelines for those reviews?
 - Is evidence of student learning part of the review?

Recommended Readings_____

The following readings are recommended along with the references cited in this chapter.

- Allen, M. J. (2006). *Assessing general education programs*. Bolton, MA: Anker.
- Calhoun, E. (1993). Action research: Three approaches. *Educational Leadership*, 51(2), 62–65.
- Ewell, P. T. (2002). *An emerging scholarship: A brief history of assessment*. In T. W. Banta & Associates, *Building a scholarship of assessment* (pp. 3–25). San Francisco: Jossey-Bass.
- Ewell, P. T. (2004). *The examined life: Assessment and the ends of general education*. Paper presented at the Association of American Colleges and Universities General Education and Assessment Conference, Long Beach, CA.

- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage.
- Lake, C., Harnes, P., Guill, D., & Crist, C. (1998). *Defining assessment*. Retrieved September 16, 2008, from www.essentialschools.org/cs/resources/view/ces_res/124.
- Leskes, A. (2002). Beyond confusion: An assessment glossary. *Peer Review*, 4(2/3), 42–43.
- Leskes, A., & Miller, R. (2005). *General education: A self-study guide for review and assessment*. Washington, DC: Association of American Colleges and Universities.
- Leskes, A., & Wright, B. (2005). *The art and science of assessing general education outcomes: A practical guide*. Washington, DC: Association of American Colleges and Universities.
- Seeley, M. M. (1994). The mismatch between assessment and grading. *Educational Leadership*, 52(2), 4–6.