CHAPTER 1

DEFINING ASSESSMENT

The concept of assessment resides in the eye of the beholder. It many definitions, so it is essential that anyone who writes or speaks about assessment defines it at the outset.

Some Definitions

In common parlance, *assessment* as applied in education describes the measurement of what an *individual* knows and can do. Over the past three decades, the term *outcomes assessment* in higher education has come to imply aggregating individual measures for the purpose of discovering *group* strengths and weaknesses that can guide improvement actions.

Some higher education scholars have focused their attention on the assessment of student learning. Linda Suskie, for instance, in the second edition of her book *Assessing Student Learning: A Common Sense Guide* (2009) tells us that for her, the term *assessment* "refers to the assessment of student learning." In the first edition of this book, we also adopted the focus on student learning:

> Assessment is the systematic collection, review, and use of information about educational programs undertaken for the purpose

of improving student learning and development. (Palomba and
Banta, 1999, p. 4)

The term *assessment* in higher education has also come to encompass
the entire process of evaluating institutional effectiveness. Reflecting her
career in applying her background in educational psychology in program
evaluation, the first author of this book uses this definition:

Assessment is the process of providing credible evidence of

- resources
- implementation actions, and
- outcomes

undertaken for the purpose of improving the effectiveness of

- instruction,
- programs, and
- services

in higher education.

In this book, the term *assessment* will certainly apply to student learning.
But we also use it to describe the evaluation of academic programs, student
support services such as advising, and even administrative services as we
look at overall institutional effectiveness.

We will describe the assessment of student learning as well as of instruc-
tional and curricular effectiveness in general education and major fields of
study. We will consider methods for assessing student learning and program
effectiveness in student services areas. We also will present approaches to
assessing student learning and program and process effectiveness at the
institutional level. In fact, the most meaningful assessment is related to
institutional mission.

Disciplinary accreditation is a form of assessing program effectiveness
in a major field. Regional accreditation is a form of assessing institutional
effectiveness. Both are powerful influences in motivating and guiding
campus approaches to assessment. Federal, state, and trustee mandates
for measures that demonstrate accountability may determine levels of per-
formance funding and also shape campus assessment responses. We will
discuss the many external factors that impel college faculty and administra-
tors to undertake assessment activities.

Our guiding principle in this book, however, will be to present approaches to assessment that are designed to help faculty and staff improve instruction, programs, and services, and thus student learning, continuously. Assessment for improvement can also be used to demonstrate accountability. Unfortunately, assessment undertaken primarily to comply with accountability mandates often does not result in campus improvements.

Pioneering in Assessment

In his book *The Self-Regarding Institution* (1984), Peter Ewell portrays the first work in outcomes assessment of three institutions. In the early 1970s, Sister Joel Reed, president of Alverno College, and Charles McClain, president of Northeast Missouri State University, determined that the assessment of student learning outcomes could be a powerful force in improving the effectiveness of their respective institutions. Alverno faculty surveyed their alumnae to find out what their graduates valued most in terms of their learning at Alverno (Loacker and Mentkowski, 1993). Survey findings shaped faculty development of eight abilities, including communication, analysis, and aesthetic responsiveness, that would become the foundation for curriculum and instruction at Alverno. In addition to work in their own discipline, Alverno faculty were asked to join cross-disciplinary faculty specializing in one of the eight core abilities. Alverno's (2011) "assessment as learning" approach has transformed that college, increasing its reputation among students and parents, its enrollment, and its visibility in the United States and abroad as a leader in conducting conscientious and mission-centric assessment.

At Northeast Missouri State University, President McClain and his chief academic officer, Darrell Krueger, became early advocates of value-added assessment, giving tests of generic skills to their freshmen and seniors and tracking the gain scores. In addition, department faculty were strongly encouraged to give their seniors an appropriate nationally normed test in their major field if one existed. McClain famously asked his department chairs one persistent question: "Are we making a difference?" meaning, "How are our students doing on those tests we're giving?" (Krueger, 1993). The early emphasis on test scores had the effect of raising the ability profile of Northeast Missouri's entering students. Subsequently the faculty and administration decided to pursue and gain approval from the state as Missouri's public liberal arts institution, with the new name of Truman State University.

The third pioneering institution profiled in Ewell's book was the University of Tennessee, Knoxville (UTK). Whereas Alverno's and Northeast Missouri's assessment initiatives were internal in their origins and aimed at improving institutional effectiveness in accordance with institutional mission, UTK was confronted with the need to address an external mandate—a performance funding program instituted in 1979 by the Tennessee Higher Education Commission and the Tennessee state legislature. Initially UTK's chancellor, Jack Reese, called the requirements to test freshmen and seniors in general education and seniors in their major field, conduct annual surveys of graduates, and accredit all accreditable programs "an abridgement of academic freedom." His administrative intern at the time, Trudy Banta, thought the performance funding components looked like elements of her chosen field, program evaluation. She took advantage of a timely opportunity to write a proposal for a grant that the Kellogg Foundation would subsequently fund: "Increasing the Use of Student Information in Decision-Making." For the first three years of addressing the external accountability mandate, faculty and administrators charted their own course on the performance funding measures on the basis of their Kellogg Project. While the amount of the Kellogg funding was tiny—just ten thousand dollars—for research-oriented faculty, the "Kellogg grant" gave them the opportunity to begin testing of students and questioning of graduates in their own way. Within five years, UTK was recognized by the National Council for Measurement in Education for outstanding practice in "using measurement technology" (Banta, 1984).

By 1985 three additional states joined Tennessee in establishing performance funding programs for their public colleges and universities. Colorado, New Jersey, and Virginia issued far less prescriptive guidelines than Tennessee, however. The state higher education organizations and legislatures in the three new entries provided examples, but left it to their public institutions to select or design tests and other measures to demonstrate their accountability.

In his 2009 paper for the newly formed National Institute for Learning Outcomes Assessment (NILOA), Ewell notes that "two decades ago, the principal actors external to colleges and universities requiring attention to assessment were state governments." However, by the 1990s, mandates in several states were no longer being enforced because of budget constraints, and so attention turned to other goals, such as higher degree completion rates. Tennessee remained an exception in continuing to employ several learning outcomes measures in its long-established performance funding program.

In 1988, Secretary of Education William Bennett issued an executive order requiring all federally approved accreditation organizations to include in their criteria for accreditation evidence of institutional outcomes (US Department of Education, 1988). During the next several years, the primary external stimulus for assessment moved from states to regional associations as they began to issue specific outcomes assessment directives for institutional accreditation, and discipline-specific bodies created such guidelines for program accreditation. The 1992 Amendments to the federal Higher Education Act (HEA) codified assessment obligations for accrediting agencies, and subsequent renewals of the HEA have continued to require accreditors to include standards specifying that student achievement and program outcomes be assessed. It has taken some accreditors longer than others to comply, however. Accreditors of health professions were in the vanguard, followed by social science professions like education, social work, and business. Engineering accreditors initiated "ABET 2000" standards in 1997 (ABET, 2013). The first trial balloon for standards related to student learning outcomes in law was launched in 2013, for approval within three years (American Bar Association, 2013).

By the time NILOA's first survey of chief academic officers was undertaken in 2009, accreditation—either disciplinary or regional, or both—was being cited as the most important reason for undertaking assessment. According to Ewell (2009), the shift in stimulus from state governments to regional accreditors had the important effect of increasing the emphasis on assessment to guide improvement in addition to demonstrating accountability. Advocating congruence of assessment and campus mission is another hallmark of the influence of accrediting agencies on outcomes assessment. A July 19, 2013, statement of Principles of Effective Assessment of Student Achievement endorsed by leaders of the six regional accrediting commissions and six national higher education associations begins, "[This] statement is intended to emphasize the need to assess effectively student achievement, and the importance of conducting such assessments in ways that are congruent with the institution's mission" (American Association of Community Colleges et al., 2013).

The pendulum is swinging once again with respect to state interest in assessment. In spring 2010, the National Center for Higher Education Management Systems surveyed state higher education executive offices concerning policies, mandates, and requirements regarding student outcomes assessment (Zis, Boeke, and Ewell, 2010). According to study results, eight states, including Minnesota, Georgia, Tennessee, and West Virginia, were unusually active in assessment, some requiring common testing. Some

states have systemwide requirements rather than state requirements. For many years, students at the campuses of the City University of New York were required to obtain a minimum score on a locally developed standardized examination in order to earn their degrees.

More recently, declining global rankings, rising tuition and student debt, and poor prospects for employment of college graduates have alarmed state and federal decision makers (Miller, 2013). This has prompted an emphasis on productivity and efficiency in higher education, which is now seen as an engine of the economy and of the nation's competitiveness (Hazelkorn, 2013). Many state reporting systems are focusing more on graduation rates, job placement, and debt-to-earnings ratios than on measures of student learning. The Voluntary Framework of Accountability for community colleges contains measures not only of how many students obtain degrees, but of how many pass remedial courses, earn academic credit, transfer to another institution, and get a job (American Association of Community Colleges, 2013). President Barack Obama's administration has proposed a College Scorecard (White House, 2013). The emphasis on producing numbers of degrees and job-ready employees has alarmed educators. They fear that educational quality will suffer if too much weight in funding regimes is placed on simply graduating more students or turning out majors who are prepared for today's jobs rather than with the abilities to adapt to ever-changing workplace demands. As Margaret Miller puts it, "The completion goal is downright pernicious if it entails the minting of an increasingly worthless currency" (2013, p. 4). In addition, emphasizing college completion in a shorter time frame could encourage institutions to raise their entrance requirements to be sure they enroll students who are best prepared for college work, which could make a college education unattainable for those who need it most.

As a result of all these external influences, as well as internal interests in obtaining guidance for continuous improvement of student learning and institutional effectiveness, increasing numbers of faculty have been called on to participate in assessment. Some assume leadership roles, serving on campuswide committees charged with planning the institution's overall approach to assessment or designing a program to assess general education. A greater number are involved at the department level, helping to design and carry out assessment of programs or courses for majors. Attendance at national, regional, state, and discipline-specific assessment conferences attests to continued interest in sharing assessment information. In fact, for more than a decade, the number of participants at the annual Assessment Institute in Indianapolis, the oldest and largest

assessment conference in the United States, has approached or exceeded one thousand. This book is designed to fill some of the continuing need for information about assessment.

Quality Assurance: An International Perspective

Interest in obtaining evidence of accountability from postsecondary institutions emerged as a worldwide phenomenon in the mid-1980s. In Europe, China, Australia, South Africa, and other countries, as in the United States, stakeholders in higher education have become increasingly concerned about the value received for resources invested, accommodating increasing numbers and diversity of students, covering cost increases with resources spread over an ever-growing array of services, developing a workforce with skills competitive in a global marketplace, and producing graduates with credentials that are transferable across cultural and national boundaries. Since postsecondary education is managed by the central government in most other countries, initial accountability-related actions were national in scope and focused on a process commonly referred to as quality assurance (QA). Self-study and external peer review, including site visits, were encouraged. Given the commanding role of central education ministries, voluntary associations of institutions like the regional accrediting agencies in the United States were virtually unknown in Europe and elsewhere prior to 1985. In that year, thirteen universities in the Netherlands formed the Association of Cooperating Universities and began a six-year cycle designed to conduct peer reviews in the same year of all universities offering degrees in a particular discipline (Vroeijenstijn, 1994). Few other countries have followed the Netherlands in developing associations of institutions to carry out QA procedures.

Prior to 1990, higher education in most other countries had been a privilege for the economically and intellectually elite: only 10 to 15 percent of the college-going age population was enrolled. In the next three decades, this percentage increased to 50 percent and higher in Europe and beyond. When only the privileged and gifted, who were generally motivated to succeed and able to navigate postsecondary education with minimal guidance, constituted the population of university students and their tuition was paid by governments, student services beyond an admissions office and a student housing staff were not needed. But with "massification," students with diverse needs necessitated the creation of a full range of student services and the institutional research function to support this infrastructure (Banta, 2013).

Although an emphasis on student success has come only recently to Europe and elsewhere, European students are much more influential in steering QA initiatives than US students are in shaping outcomes assessment. At the University of Freiburg in Germany, undergraduates in psychology made their own list of learning outcomes and presented it to the psychology faculty (Banta, 2009). In the United Kingdom, the National Student Survey is administered at all universities. The results, reported publicly, have become an important factor in judging the quality of institutions. There is now a European Students' Union (ESU, 2014) that voices its concerns about quality assessment on occasion. The ESU is an umbrella organization for forty-seven national unions of students and has headquarters in Brussels, from whence its officers can lobby the European Union.

Policies of the European Union have created transparent borders and a common currency that combine to encourage workers to migrate from one country to another. By 1999, it was clear that some standardization was needed in order to make sense of university degrees produced by programs differing in content, length, and curricular structure. In that year education ministers representing twenty-nine European countries met in Italy and drew up the Bologna Declaration, which called for comparable degrees based on a common framework of degree levels, a credit-based system that would facilitate international transfer, and a cooperative system of quality assurance (Ewell, 2004). In 2000 the European Network for Quality Assurance in Higher Education (ENQA) was established to promote European cooperation in ensuring quality. Australia, New Zealand, and Hong Kong are among the many other countries that also have national QA organizations that conduct institutional audits and disseminate information about good practice.

In 2000 ENQA initiated a process designed to develop field-specific reference points, including learning outcomes, that could guide students to a credential in a field of study. The process is called "tuning" to reflect its "attempt to steer a course between identical cross-national standards and institutional autonomy with respect to degree standards" (Ewell, 2004, p. 12). Tuning in Europe has inspired similar efforts in Latin America and Russia, as well as the Lumina Foundation's Tuning USA project involving several disciplines in Indiana, Kentucky, Minnesota, Texas, and Utah (Adelman et al., 2014, p. 38).

Faculty development has been a hallmark of QA initiatives, particularly in Britain. But that has not kept British academics from voicing their resentment toward QA because it is imposed by external

authorities (Banta, 2009). In addition, the tension between assessment for improvement and assessment to address accountability demands is as palpable in Europe and elsewhere as it is in the United States (Harvey and Williams, 2010).

Prior to 2000, most of the scholarship in postsecondary outcomes assessment in other countries emphasized national and state (e.g., *landers* in Germany) quality assurance initiatives, while in the United States, this emphasis was on classroom assessment and assessment in general education and academic disciplines (Banta, 2000). Since 2000, assessment scholarship in all countries has also included a focus on institutional effectiveness: electronic institutional portfolios, quality management (e.g., Total Quality Management and ISO 9000), academic audits (external peer review), performance indicators (e.g., key performance indicators), league tables or rankings (e.g., *U.S. News & World Report* rankings and U-Multirank in Europe), and surveys for students (e.g., the National Survey of Student Engagement, the Australian Survey of Student Engagement), faculty, graduates, and employers.

Assessment Purposes

Assessment is more than the collection of data. To make assessment work, educators must be purposeful about what they collect. As a basis for data gathering, they must clarify their goals and objectives for student learning and be aware of where these goals and objectives are addressed in the curriculum. After data are gathered, educators must examine and use assessment results to improve educational programs.

Alexander Astin (1985, 1991) helped refocus thinking about the quality of higher education from resources such as library holdings and student SAT scores toward educational outputs such as knowledge, skills, and values. He also argued that the results of education must be understood in relationship to both inputs and environment. Knowing where students wind up is only part of the story; information about where they start and what they encounter along the way is also necessary. Without this context, it is difficult to use outcome results for improvement. Questions about student experiences need to be considered in planning assessment programs. What kinds of courses do students take? What opportunities are available for out-of-class learning and development?

The overriding purpose of outcomes assessment is to understand how educational programs are working and to determine whether they

are contributing to student growth and development. Hence, the ultimate emphasis of outcomes assessment is on programs rather than individual students. At its most useful, assessment provides information about students as a group—information that can be aggregated across sections of a single course and is meaningful across courses. Assessment indicates what the experiences of students add up to and what these experiences imply about educational programs. It enables educators to examine whether the curriculum makes sense in its entirety and whether students, as a result of all their experiences, have the knowledge, skills, and values that graduates should possess. Program assessment helps determine whether students can integrate learning from individual courses into a coherent whole. Interest is focused on the cumulative effects of the educational process. Assessment helps us look at programs in a holistic way.

Assessment participants on most campuses recognize that assessment contributes to improved teaching and learning and that it also satisfies reporting requirements, but it serves other purposes as well. For some institutions, assessment's greatest benefit is fostering academic introspection—making the institution more self-conscious about what its programs are accomplishing. In 2004, Calvin College faculty redesigned their assessment program after receiving a negative review from an accreditation team. A first step was to articulate a philosophy of assessment that was compatible with institutional culture. Assessment is strongly affirmed and framed within the concept of reflection. Everyone is encouraged to be reflective, not just as an individual but as a member of a department and the community as a whole (Bradley, 2009).

St. John's University (2014) leaders conceptualize assessment as a three-way conversation among students, faculty, and administration. It is a continuing dialogue in which all participants learn from each other. At Carnegie Mellon University (n.d.a), assessment information is intended to help instructors "refine their teaching practices and grow as educators."

Other recognized benefits of assessment include attracting prospective students, improving institutional image, and developing funding opportunities. Montgomery College (n.d.) leaders see assessment as providing a way for faculty and administrators to tell their story to others, including politicians, employers, and potential donors. And at the Massachusetts College of Liberal Arts (2013), faculty find that they can use assessment information to inform planning, document needs, assist with grant writing, and improve each academic program and its standing in its discipline.

The purposes a particular university or college chooses for assessment may be captured in an assessment definition. Occasionally campuses use an existing definition, such as the ones we offered at the start of this chapter. More often, they develop their own unique definitions. Baruch College faculty (2014) describe assessment as an important way to make informed decisions and a means to pursue excellence in teaching and learning.

Values and Guiding Principles

In addition to a definition of assessment, many campuses develop statements of values or guidelines for carrying out assessment, as well as lists of best practices. Examining such documents reveals a great deal about the particular campus approach to assessment.

Some statements emphasize collaboration and inclusion in developing assessment initiatives, and many affirm that faculty are the individuals best suited to carry out assessment in their disciplines. Many also specify that assessment results will be used to evaluate academic programs, activities, and student services but not individual faculty or staff.

Issues of timing and methods also are addressed. Montgomery College (n.d.) educators value an assessment approach that is "as simple and manageable as possible. The process cannot become so onerous that it hampers or interferes with the delivery of the educational experience that it attempts to assess and improve." College leaders also expect faculty and staff to use assessment results. If assessment is primarily a reporting tool, "then this effort will have been deemed a failure."

St. Ambrose University (2014) faculty value assessment efforts that are timely, efficient, and feasible; use existing data and instruments if possible; and are informed by scholarship and good practice. Hartwick College (2014a) educators value assessing "smarter not harder" by using methods and strategies that are simple, sustainable, and iterative and inform efforts to improve teaching and learning.

Along with, or in place of, a statement of values, some educators find it helpful to develop a set of operating principles or guidelines that clarify the purposes and intended uses of assessment information. Guiding principles at North Carolina State University (2006) call for curiosity and intellectual dialogue, involvement that is recognized and rewarded, and relevant evidence that is used in many ways, including highlighting excellence and informing planning and resource management. At Cabrillo College (n.d.),

faculty and staff view assessment as a way to identify issues that can benefit from campuswide discussion.

Rather than guiding principles that often set specific expectations for a particular campus, lists of good practices may present ideals to which campuses can aspire. For example, the "Characteristics of an Effective Assessment Program" statement shared on the James Madison University (n.d.) website calls for evidence that resources are moved or reallocated in response to assessment results and that traditional and technological delivery modes are compared with respect to student learning and development.

Practitioners on many campuses find it useful to prepare a set of frequently asked questions for faculty, staff, and students. In addition to answering questions about the who, what, and why of assessment, FAQs can deal directly with issues that are problematic. Most lists answer the question of how faculty grades and assessment information differ from each other. As we elaborate in Chapter 3, grades reflect participation, effort, and attendance, as well as learning. And grades alone generally fail to capture enough information about mastery of learning objectives to provide direction for improvement.

Other frequently asked questions may reflect specific campus interests. For example, the set of FAQs available on Stanford University's (2014) website asks: "What is the difference between assessment and research?" The answer explains that assessment is a type of action research with the primary goal of improving practice rather than generating theoretical knowledge. Experimental control and random assignment are often not possible in instructional settings. Nyack College (2014) FAQs address the practical question of whether all faculty teaching the same course will be required to use the same textbook as the result of participating in program assessment. The answer is, "Not necessarily." It is the skills and knowledge to be assessed that are common, not the textbook.

On some campuses, the greatest challenge in implementing department-based, faculty-involved assessment is to develop a common language about the meaning of and uses for student assessment. Faculty are used to thinking about assessment of individual student learning and dealing with curriculum issues; they are much less familiar with using assessment for the purpose of improving programs. Thus, on a number of campuses, sets of definitions or glossaries for faculty and staff have been developed. The Internet Resources for Higher Education Outcomes Assessment site maintained at North Carolina State University provides links to nearly

fifty campus glossaries. The web page for California State University Long Beach (n.d.b) contains some two dozen assessment definitions.

All of these materials should reflect the values and interests of the many stakeholders in assessment, not just of a few decision makers. Formal and informal discussions about the aims of assessment need to be inclusive if the campus is to embrace assessment as a way to provide evidence to guide improvement.

As we have illustrated, outcomes assessment is a many-faceted concept. When one uses the term, a definition must follow to ensure that the ensuing conversation will take place in the appropriate context. Statements of values and guiding principles are important to the clarity of this conversation.