3

STUDENT LEARNING OUTCOMES ASSESSMENT AT THE PROGRAM AND INSTITUTIONAL LEVELS

Thomas Judd and Bruce Keith

In the widely and often quoted book, *Academically Adrift: Limited Learning on College Campuses*, Richard Arum and Josipa Roksa (2010) conclude that college students “are failing to develop the higher-order cognitive skills that it is widely assumed college students should master. These findings are sobering and should be a cause for concern” (p. 121). Is this true? Are our colleges and universities failing in their mission? If so, why should an institution put scarce resources into assessing student learning outcomes when there are many other issues vying for attention? The answer to these questions lies in the very purpose of assessing student learning outcomes: using evidence for determining if students are achieving the desired outcomes, and for improving instruction based on the findings of assessment.

In this chapter, we will provide an overview of the background, purpose, context, and methodology of a system to assess the student learning outcomes of a college at the institutional and program levels. Assessment of student learning outcomes not only provides evidence regarding an institution’s achievement of the mission and improves the delivery of instruction, it also informs strategic planning and facilitates change and development at the course, program and institutional levels.

We begin with a brief background of the development of assessment initiatives, followed by a discussion of mission statements and the need for alignment with student outcomes and educational activities, methodologies for effective student outcome assessment, sources of evidence that provide useful and actionable information, issues in the sustainability of an assessment system, and conclude with a case study of assessing general education student learning outcomes.

**Context for Outcomes Assessment**

For several decades researchers have been engaged in a wide variety of approaches to assess student learning outcomes: the education-related consequences of students’ postsecondary educational experiences (U.S. Department of Education, 1997). In fact, *expansive* is the word of choice used by Pascarella and Terenzini (2005) to characterize the changes they observed in the previous decade of research on the impact of a college experience on students. Their comprehensive review documented an expansion in the sheer volume of studies, the types of institutions studied, the underlying developmental framework, the variety and influences of policy concerns and the diversity of research approaches.

With such a potentially bewildering array of purposes, contexts and methods for assessing outcomes, it is conceivable that the function and practice of assessment at an institution can become lost
in complexity. It is no wonder that many assessment programs have difficulty in gaining traction, lose their focus, have little faculty buy-in, and subsequently wither on the vine for lack of effectiveness, usefulness or relevance. Indeed, several authors have taken a practical step-by-step approach in their texts to provide institutions with excellent guidance as they take their first steps into process of assessment of student learning outcomes (e.g., Leskes and Wright, 2005; Suskie, 2004; Walvoord, 2009). These authors describe a process that (a) defines the student learning outcomes, (b) articulates an alignment of the student outcomes with curriculum activities, (c) collects evidence relevant to the student outcomes, and (d) uses the results to improve student learning.

**Background: Assessment Initiatives**

The assessment movement gained momentum from calls for accountability in higher education at a time when there was little data available that reliably and accurately told a comprehensive story of the effects of college on students (Ewell, 1984, 2002). Ewell (2002) traces three national movements that influenced the debate over the efficacy and form of assessment, and influenced the scholarship of assessment: Value-Added, Total Quality Management (TQM), and the National Educational Goals Panel (NEGP). These movements focused debate on whether assessment should address student attainment or institutional contribution (Value-Added), on the need to address systems and processes (TQM), and on the notion of a national accountability through standard outcomes assessment at the college level (NGEP).

Partially in response to the calls for accountability, institutional accreditation has moved from an examination of fiscal solvency, faculty and administrative credentials and vitality of resources to an examination of documented institutional effectiveness and student outcomes. The evidence required for successful accreditation is focused on student outcomes and effectiveness data that is collected “within the context of the mission of the institution, the suitability and effectiveness of processes designed to accomplish institutional goals, and the institution’s ability to fulfill its purposes” (Council of Regional Accrediting Commissions, 2003, p. 2).

In the broadest context, outcomes assessment is an integral component of a comprehensive planning process in which student outcomes are part of a system that encompasses the assessment of the effectiveness for the entire institution. Within a given context, each functional component of an institution has a mission and set of corresponding goals that align with those of the institution. Both student learning outcomes and institutional effectiveness data are collected to inform the institution about how successful it is in achieving its mission and to fuel a comprehensive strategic planning process (Middaugh, 2009). Therefore, the assessment of student learning outcomes is primarily about improving student learning through a systematic, intentional, and local process.

**Institutional Mission and Alignment With Goals, Learning System and Student Outcomes**

The mission and goals . . . are used to develop and shape its programs and practices and to evaluate its effectiveness (Middle States Commission on Higher Education [MSCHE], 2006, p. ix). The starting point for an assessment process should be a well crafted and meaningful mission statement that is embraced by all constituents of the institution. The mission provides guidance for an institution’s goals and learning outcomes; the curriculum provides a framework for the process and activities that ensure the educational environment will lead to achievement of the student learning outcomes.

In one accrediting agency’s definition an institution’s mission “clearly defines its purpose within the context of higher education and indicates who the institution serves and what it intends to accomplish . . . [Mission statements] are developed and recognized by the institution with the participation of its members and its governing body” (MSCHE, 2006, p. ix).
The alignment of the goals with the mission statement is an indication that the institution or program has a clear sense of what it is about, and what the expectations for its program completers are. Again, in the words of an accrediting agency, “The institution’s stated goals, consistent with the aspirations and expectations of higher education, clearly specify how the institution will fulfill its mission” (MSCHE, 2006, p. ix).

The mission and goals or student outcomes give us two points: the institution’s or program’s stated purpose at one end, and the anticipated result (goals or outcomes) of the experiences that occur during the program at the other. All that happens between these two points are what we refer to as the curriculum. In this sense, the curriculum may include experiences beyond the courses and classroom that are intentionally arranged to contribute to the outcomes. The activities that comprise the curriculum are driven by the outcomes—one should be able to discern an alignment between these activities and the outcomes to which they are contributing.

The alignment between the mission statement, the student learning outcomes, and the design of learning activities cannot be understated; the purpose of an education, as defined by the student learning outcomes, is wholly dependent on the design of the learning system (Spady & Schwahn, 2010). For example, an institution that wishes its graduates to demonstrate an increased capacity for civic responsibility should be able to show that there are activities designed to contribute to the development of that outcome, such as a service learning component, built into the context of every major.

Ancillary activities—first year experience, internships, study abroad, clubs, sports, immersion, multi/interdisciplinary, service learning, cultural events—are but a few of the activities that may contribute to general education outcomes and should be considered in the learning system design.

An important step in the assessment of student learning outcomes is the identification of places in the curriculum where activities aligned with the outcomes occur. For example, constructing a matrix of student learning outcomes and the content of course syllabi can identify locations in the curriculum where student learning outcomes are being addressed, as well as potential places for gathering direct evidence of student performance of the outcomes.

### Institutional Level Student Outcomes: General Education

General Education student outcomes are the knowledge, skills and competencies all graduates of the institution can be expected to demonstrate, regardless of the major. General education outcomes can be intellectual, social, civic, political, religious, moral/ethical, physical, or occupational. They can be achieved through a series of specific courses, infused throughout the curriculum across disciplines, reinforced in the major, or through a combination of strategies.

Although individual institutions or university systems determine their own version of general education outcomes, there are several national models that can provide guidance. The Association of American Colleges and Universities (AAC&U, 2007) describes seven areas for general education: written and oral communication, critical thinking and research skills, quantitative reasoning, information literacy, ethical reasoning, civic engagement, intercultural skills, application of learning, and integration of learning. AAC&U has also published a comprehensive set of rubrics for each of these outcome areas that have been developed by panels of curriculum and higher education experts. The Wabash National Study of Liberal Arts Education (2006–2009; Center of Inquiry in the Liberal Arts, n.d.), a project with 49 member institutions and 17,000 students, is examining seven liberal arts outcomes: effective reasoning and problem solving, inclination to inquire and lifelong learning, integration of learning, intercultural effectiveness, leadership, moral reasoning, and well-being.

National accrediting agencies also offer their views on appropriate general education learning outcome areas. MSCHE, for example, asks its member institution to have general education student learning outcomes in the areas of written and oral communication, scientific reasoning, quantitative
reasoning, critical analysis, technological competency, values and ethics, intellectual growth, and cultural and global awareness.

The commonality of these student outcome areas at the institutional level can be viewed through several categorization systems with an eye towards comprehensiveness. Astin (1993), for example, describes a taxonomy with two dimensions: psychological/behavioral and cognitive/affective that, when combined, produce a 2×2 taxonomy of student outcomes. Ewell (1984) categorizes general education outcomes into areas of knowledge, skills, attitudes, and relationships.

While the knowledge, skills and competencies of general education are ubiquitous in higher education, there are clear indications that general education goals are moving into other more challenging affective spheres. Spady and Schwahn (2010) take a broad view of the role colleges play in preparing graduates for life by suggesting that institutions should consider eight dimensions, or spheres of living, for development of students including: personal potential and wellness, learning challenges and resources, life and resource management, close and significant relationships, group and community membership, work and productive endeavors, physical and cultural environment, and purposeful and fulfilling pursuits. W. Astin, H. Astin, and Lindholm (2010) describe the role colleges can play in developing the inner, spiritual, and affective lives of students as they develop during the college years.

### Program or Majors Level Student Learning Outcomes

Program-level outcomes represent student learning outcomes for a major or discipline. They may be derived from the collective expertise of the faculty, and can be informed or directed by local needs, career requirements, and professional or accreditation standards. Program-level student learning outcomes may also provide an opportunity for reinforcement and infusion of general education outcomes.

Engineering and technology programs, for example, have to meet the standards determined by Accreditation Board for Engineering and Technologies (ABET)—the accrediting body for engineering and technology programs—which requires a set of 14 student learning outcomes in each of its accredited programs. Majors without accrediting bodies may have professional associations with recommended student learning outcomes that can guide their development for programs. With or without external sources to guide or dictate student learning outcomes, faculty in a program will need to determine the full range of student learning outcomes and standards of performance for their programs.

Institutional and program-level outcomes may look very similar, but a full set of program-level goals should have the specificity necessary to distinguish the distinctive learning outcomes of a given major. For example, there may be a quantitative reasoning student learning outcome at the institutional level, while the psychology program may have a student learning outcome for statistical analysis in a research methodology outcome. As with general education student learning outcomes, program student learning outcomes should be aligned with the program mission, the general education student learning outcomes, and the mission of the institution.

### Levels of Performance/Development

While clear statements of expected student learning outcomes that are aligned with the mission are a key component of assessment, the level of performance expected is another issue that can be informed with reference to standards generated internally or externally. The standard for completing the mile run can be based on a norm for gender and age groups, referenced to personal improvement, or referenced to minimum time standard set by knowledgeable faculty. While this standard may be relatively straightforward, when the outcomes fall into some of the more complex and
difficult to measure categories, the development of standards becomes increasingly challenging. When a college says that their graduates will be able to understand global issues and perspectives, or clearly communicate ideas in written and oral form, how is a standard of performance determined, and at what level? While there is some subjectivity inherent in determining a standard of performance, there are methods of clarifying aspects of the standards.

With the cognitive, affective and psychomotor domains, for example, Bloom’s (1956) Taxonomy has proven to be a powerful resource for educators developing objectives that incorporate a specified level of performance in the standard. The cognitive domain, as described by Bloom has six levels of student mastery of material in this taxonomy, starting from knowledge at the simplest level, through comprehension, application, analysis, and synthesis to evaluation at the most complex. Lists of active verbs that serve as keys for performance at each of the levels abound in many sources.

Once the potential levels of performance are determined, the level that will be used as a basis for a measure of success may be specified as a standard. If a general education communication student learning outcome is assessed by an institution-wide standardized test of reading and writing skills which all students must pass prior to graduation, the standards outcomes and standards are clear, direct, easy to understand, and easy to determine if the college is achieving this goal of the mission. (See Pitoniak and Morgan, this volume, for a discussion of what constitutes passing.)

This is not the case with all outcomes. The goal to understand global issues and perspectives raises the question: what does it mean to understand, and how do we state it in such a way that it can be measured?

In Bloom’s view, understanding is a part of the second level of cognitive activities he terms comprehension. He also provides us with verbs that describe actions that students can use to demonstrate their understanding: comprehends, converts, defends, distinguishes, estimates, explains, extends, generalizes, gives examples, infers, interprets, paraphrases, predicts, rewrites, summarizes, translates. These verbs can be used to define outcomes with a mind for creating assessment measures. Understanding global issues, for example, may then be defined as summarizing how scarce energy resources (or any other global issue) may affect international relations. An outcome embracing a higher level of cognitive skill may state that graduates evaluate a national policy in an international context. This could require that students compare, explain, critique, and draw conclusions about an energy policy in a global context.

The level of development that students reach through their college experiences can also be examined through one of the theories of student development and used as a basis for developing standards for assessment. Evans, Forney, Guido, Patton, & Renn (2009), Robert Kegan (1982), William Perry (1970), and Arthur Chickering and Linda Reisser (1993), for example, each describe models of developmental changes of students through the young adulthood years of college. Typically, the models show that as individuals gain experience and confidence, external controls give way to internal controls, accompanied by expanded horizons and movement from self-centered to interpersonal sensitivity, from simplicity to complexity, from dependence to independence, from impulsivity to self-control, and from instinctual to principled action. By developing student learning outcomes and standards that use a developmental modal as a basis, realistic expectations for levels of performance and the sequencing of developmental activities can be consistently defined.

Methodology: Does Assessment of Student Learning Outcomes Require Scientific Rigor?

Reliability, validity, sampling, and item construction are all components of scientific rigor, and are topics addressed elsewhere in this volume. The question for assessment is how rigorously should these concepts apply to a system of student outcomes assessment? The prime function of assessment...
of student outcomes is for local decision making, and does not necessarily require the full range of rigor associated with experiments designed to support cause and effect conclusions.

The notion that assessment implies a judgment also implies that the judgment is made with confidence in the supporting evidence. Judging whether an outcome has been achieved can seem quite arbitrary if the bar for success is set without reference to any level of performance that is meaningful, or with a methodology that is flawed or renders the evidence difficult or impossible to interpret. As described in a previous section, such judgments can be made from a set of standards, from peer comparisons or from documenting personal or individual growth. Standards can be derived from external sources such as professional consensus, accrediting bodies, industry needs assessments, or from agreed upon local standards. Peer comparisons, more commonly referred to as benchmarks (see Seybert, Weed, & Bers in this volume), address how well one group does in comparison to a set of their institutional, program or departmental peers. Personal growth as a point of reference depends less on how well a group does in comparison to peers or external standards, than on how much individual students have developed from their starting point at the institution.

These three contexts for judgment—standards, benchmarks and personal growth—have their appropriate place for various kinds of decision-based outcomes. For example, if an outcome requires students to demonstrate the ability to run a mile, peer comparisons and personal growth are irrelevant. The standard is completing the run, regardless of the performance of peers, or whether the student is a recruited track athlete or had never run more than a few meters prior to admission to the institution. If the outcome is to produce the best nurses in the field, benchmark comparisons of outcome measures fit the bill. On the other hand, if the outcome is to develop in students an increased capacity for civic responsibility, personal growth may be the most appropriate context for judgment. The phrase ‘to develop’ suggests that the institution is interested not only in where the student is at the end of the educational experience, but how much progress has been made since the student entered the institution—a central feature of a ‘value-added’ approach. These three types of perspectives require somewhat different sets of standards.

Frameworks for Assessment

Judgments derived from student outcome data can be made with more confidence if the methodology supports judgments of causality. Anyone steeped in research methodology can recognize designs of studies that support cause-and-effect conclusions about the agents of documented change and those that show change, but cannot attribute change to a particular agent. If outcome data show that graduates have a high level of social responsibility, can one conclude that it is due to the experiences at their college rather than characteristics of students admitted or drawn to the school? Outcome data alone may not be sufficient to support that conclusion, but if it is collected in the context of a framework of other knowledge that supports causal inferences, then it may be justified. True experimental research designs take into account all possible influences on student learning outcomes and systematically vary them to identify agents of cause. It does not seem reasonable to expect all faculty or staff from diverse disciplines involved in assessment to become facile with research designs, but it is feasible to present a conceptual framework for understanding the limits of inference that parallels the logic of experimental design.

Astin (1993) described one such framework for interpreting assessment studies that takes into account inputs, environment, and outcomes, which he calls the I-E-O model. Inputs are those qualities, traits, skills, attitudes, or beliefs that students bring at their entry into the program or institution being assessed. Environment is everything that is experienced during the course of the program, and outcomes are the evidence of student performance. Inputs and environment can independently and in collusion effect the outcomes. It is not difficult to see that entering students at a highly selective
Student Learning Outcomes Assessment at the Program and Institutional Levels

private college with a mean SAT score in the highest five percent is not starting at the same place as an open enrollment institution in an inner city with a low average academic performance of entering students. Inputs clearly could impact the outcome measures in the myriad ways in which entering students differ, making judgments about the comparative effectiveness of a program unclear. Even a comprehensive description of the educational experiences could not allow a claim for their contribution to the outcomes independent of the inputs.

Not only do environments differ between institutions, departments and faculty, they also can have different effects on students with differing input characteristics, resulting in outcomes that have multiple and compounded influences.

Perhaps the most useful aspect of Astin’s I-E-O framework is a clarification of the kinds of inferences that can be drawn about the relationship between outcomes and differing combinations of the associated input and environmental components. The emphasis on outcome data by the accreditation and accountability movements resembles the outcome-only assessment model in this framework. Standardized tests and surveys given to seniors, while providing comparability to a broader population of seniors, fall into this category if they are the only source of information related to outcomes. While the outcome-only assessment model does focus on data that provides information about the success of the institution achieving its goals, it does little for identifying causal relationships between the content of programs and results. So, did those graduates with a high level of social responsibility arrive at the institution with those inclinations? Did the school contribute to the outcome, or draw students who were already on the road to social responsibility by virtue of their family and religious background and their fit with the institution’s mission? Outcome-only approaches don’t provide insight into these questions, but they do address the issue of whether graduates meet the standard set for that institutional student outcome.

The environment-outcomes approach resembles the normative approach to grading: comparisons are made between institution or programs without attention to differences at input. Comparing the written communication skills of physics majors to philosophy and music majors who all went through the same general education courses in English composition and literature may show differences between them, but ignores the fact that the students drawn to these diverse majors may have entered with different levels of skills, motivation and interests that could have influenced their written communication skill development.

The environment-only approach is typical of earlier accreditation standards, where the focus is on the program elements such as faculty qualifications, instructional materials, seat time, work load, textbooks, laboratory work, and facilities. While these factors may contribute to the ability of an institution to meet its mission, there is no information that tells us how well the mission is met, or if the environment had an impact.

Astin’s model not only points us in the direction of the kinds of conclusion we can make from the outcome data at hand, but at the kinds of evidence an assessment process may strive to collect in the face of the kinds of conclusions the process is designed to draw. Extraordinary claims require extraordinary evidence. While the claims we make in assessment may not meet the criteria of ‘extraordinary,’ they certainly require appropriate evidence. In Astin’s framework, an assessment that takes inputs and environment as well as student outcomes into account can provide the appropriate evidence to substantiate the claim for the effect of the program on the student outcome. Such claims for the causes of outcomes are not always required in the accreditation process, but we need to be cognizant of the limitations on interpreting the data that is produced.

Astin’s model is one of several frameworks for assessment. For example, a National Center for Education Statistics (NCES, 2002) review of competency-based assessment arrays outcomes in a hierarchy, with traits and characteristics as a foundation, and with skills, abilities and knowledge developed in the learning process which are then acquired as competencies. At the peak of the hierarchy are demonstrations of the competencies, and the point for assessment of
Sources of Evidence

“I give my students grades. That is evidence of their achievement. Why do we need to collect other assessment data?” This common comment from a faculty member is understandable. Grades have a long history of providing evidence of faculty evaluation of student achievement, and do represent how individual students perform. But in the context of assessment, grades alone do not typically provide evidence that allows specific curricular decisions to be made based on the results of student learning outcome assessment. Grades represent subjective evaluation of multiple aspects of a course and an individual’s predisposition toward a course and his or her external environment.

Assessment, to be useful, needs a collection of evidence that allows judgments to be made regarding the achievement of specific student outcomes. When the evidence points to a weakness in the achievement of an outcome, it should also inform the stakeholders where action can be taken to strengthen the outcome. Grades give a general sense of the faculty member’s evaluation, but what can a class average of 74% on a course, final, test, paper or project tell us about how to improve outcome performance? Certainly, a well-crafted product may do just that if the component parts are structured to align with specific outcome measures and grades for the component pieces are presented as the evidence. Grades for the component parts related to the student learning outcomes can then be summarized across students and even courses to provide direct evidence of student achievement of the outcomes. Another method for examining the relevant components of a test given to students or a student project is the use of rubrics as an ancillary tool for assessment of these products, independently or in conjunction with the grading process (see Yen and Hynes in this volume).

Not all evidence collected for assessment is equally persuasive. Indeed, as Yen and Hynes note, as calls for accountability through assessment increases among the accrediting agencies, the likelihood that they will require evidence with higher levels of reliability and validity is also increased. Conveniently, Sections 2, 3, and 4 in this volume address these issues in some detail. We will continue with the distinction between indirect and direct sources of evidence and provide some examples.

Direct and Indirect Indicators

Direct evidence is “tangible, visible, self-explanatory evidence of what students have and haven’t learned” (Suskie, 2004, p. 95). This evidence is based on performance of students. In contrast, indirect evidence provides an indication of learning, but lacks the clarity of direct evidence in observing actual performance. Carefully constructed surveys which align with student outcomes rely on the judgments of performance by the responders rather than actual measures of performance. Surveys requiring respondents to judge competence, skill level, or confidence in performing student learning outcomes are providing subjective judgments rather than demonstrations of performance. These are examples of indirect evidence, which provide one perspective of achievement of the student learning outcomes, but do not represent actual performance. Direct indicators are demonstrations of student performance on course products—such as tests, presentations, papers, portfolios or projects—that give direct evidence of student achievement of learning outcomes.

Embedded and Add-On Indicators

Higher education institutions routinely collect vast amounts of data, from administrative data to final exams, papers, projects, and grades on daily work. Not all of this is relevant for outcomes, but in this myriad collection there is a treasure trove that can be mined for assessment purposes. By identifying performance. This hierarchy aligns similar components of Astin’s framework within a different conceptual framework.
those products that relate directly to student outcomes and planning for their collection, assessment can become relatively efficient and less burdensome for faculty. By constructing the matrices of student learning outcomes and course content noted earlier, potential locations for student products can be identified. The next step is to identify the existing products that can be used for assessment. If a psychology exam in a core course asks students to apply an understanding of the development of moral decision making, faculty can compile that evidence for assessment of a student learning outcome in moral decision making. Data collected from embedded indicators aligned with the outcome statements, can grow into a body of evidence that informs outcomes assessment with little additional effort by faculty and that is clearly tied closely to the curriculum.

Embedded indicators may not always be sufficient or available for the assessment process. Although the data from embedded indicators is convenient to collect, they do not provide the opportunity for comparisons with other relevant groups. If there is a need for these comparisons, assessment may require an externally developed or nationally normed instrument such as the Defining Issues Test (Rest, Narvaez, Thoma, & Bebeau, 1999), which examines moral reasoning based on the developmental model of Kohlberg (1984), and gives national comparative data for college freshmen and seniors.

There are many nationally normed instruments available for a variety of student learning outcomes. For example, the College Learning Assessment (CLA), cited by Arum and Roksa (2010) in their assessment of the state of higher education, measures students’ ability to think carefully, reason analytically, solve problems, and communicate. In determining if a given instrument is appropriate, care must be taken in examining whether the skills being evaluated align with the student outcomes of the institution or program. If an institution’s student learning outcomes include the ability to think carefully, reason analytically, solve problems and communicate, and there is a desire for a comparison with other colleges, the CLA is an instrument that could be considered.

**Triangulation Principles**

Direct and indirect evidence that is collected for the assessment of student achievement of outcomes can vary greatly in how confident we can be that it is a valid measure of the outcome. Certainly many direct measures can be interpreted fairly confidently. Timing how fast a student can run a kilometer, measuring whether a model bridge can support a fifty pound weight, or judging the construction of a logical argument can all be deemed to be accurate measures of specific outcomes. But there are many outcomes where the use of indirect measures may be the best available or most convenient evidence. For instance, when assessing an outcome in the moral domain one could use an instrument such as the Defining Issues Test (Rest et al., 1999) to assess the level of moral reasoning ability, but this would not necessarily inform the assessment regarding moral decisions in real life circumstances. Often indirect indicators must be relied upon when direct measures are not practical. One approach is to ask graduates’ supervisors in a survey their level of confidence in the graduates’ decision making in situations that have a moral or ethical component. Although this is an indirect indicator and relies on subjective judgment, it does provide another data point in assessing that outcome from a source that has intimate knowledge of the performance of the graduate.

Assessing affective psychological or behavioral outcomes, in particular, poses problems in methodology and data collection that may be beyond the resources or expertise available at the institution. In these cases, indirect evidence such as survey data may be the most reasonable data accessible to the assessors. In spite of the rigorous methodology that has been developed in constructing, administering, analyzing, and interpreting survey data, there are still limitations in how much confidence we can have in concluding that responses on a survey reflect the actual abilities or outcomes. If a graduate indicates on a survey their confidence in their ability to write a coherent letter of recommendation supporting a position, how valid is that assertion? A student as a freshman and again as a
senior may report a high level of confidence in their written communication skill. However, their reference point for quality written communication, applied to their own abilities, may have changed greatly over the four years. While their confidence in written communication may have remained high over that time, their skill level may have improved dramatically. Interpretations of the survey data must be tempered in some way to account for the subjective nature of the information.

This source of data, while not ideal, can still contribute to our assessment of a given outcome through the method of triangulation, in which multiple sources of evidence used together measure a construct in corroboration. If the graduate makes the assertion, the graduate’s employer on a survey is comfortable with the graduate’s writing, and the data from a senior thesis provides evidence of adequate writing, the three sources of evidence converge and allow the conclusion that the writing outcome indicates that the graduate has met the standard. This is the optimum situation, where all sources converge on a single conclusion. This may not always be the case, as triangulation may also result in inconsistency or contradiction among the varied data sources (Mathison, 1988). Mathison suggests that in all three cases—convergence, inconsistency and contradiction—assessment “requires embedding the empirical data . . . with a holistic understanding of the specific situation and general background knowledge” about the broader context that the researcher brings to the assessment (Mathison, p.17). By using thoughtful triangulation in assessment we can strengthen our confidence in our conclusions and recommendations, and recognize and examine the reasons for inconsistencies and contradictions.

**Sampling**

Assessing student learning outcomes does have costs in faculty time and survey and instrument administration. It may be cost-prohibitive and even impractical to collect data from every student, every year, for every student learning outcome, especially in large institutions; and it is not necessary. Sampling principles can be applied to get a representative group of students that gives sufficient confidence in drawing conclusions from the data. Suskie (2004) suggests that small sample sizes can make the assessment process feasible by reducing the data collection burden, but the increase in sampling error should be taken into account when interpreting the results. She includes practical suggestions for determining sample sizes and interpreting from results samples. Confidence in the accuracy of the conclusions can be bolstered by using triangulation principles and by looking for consistent results among diverse samples. Sampling techniques can be applied not only to the number of student products observed, but also to the student learning outcomes. Over successive years, data can be systematically collected from different student learning outcomes so that, in the long term, every general education and program-level student learning outcome has a body of assessment evidence.

**Using and Sustaining a Process of Assessment: Faculty, Administration, and Staff Buy-In and Support**

A successful student learning outcome assessment process is embedded in the culture of the institution, and is valued by the institution and supported by administration, and is useful. Key to creating a successful process is a method for reporting and responding to the evidence. While this can be done informally, and indeed, faculty collecting assessment data often revise the curriculum based on their findings, a systematic and documented reporting process can add assurance that the assessment process has validity. A report should summarize data, draw conclusions based on the evidence, and propose recommendations for curriculum and future assessments. It is crucial that leadership acknowledge and respond in a reasonable manner to the recommendations made in the report. (See Banta and Pike, this volume.)
Indeed, the ultimate purpose and success of a system that assesses student learning outcomes is determined by how it improves student learning, and this is where faculty are most likely to see assessment’s usefulness.

Assessing General Education Student Learning Outcomes: A Case Study

The mission of the United States Military Academy (USMA) is “To educate, train, and inspire the Corps of Cadets so that each graduate is a commissioned leader of character committed to the values of Duty, Honor, Country and prepared for a career of professional excellence and service to the Nation as an officer in the United States Army” (USMA, 2007, p. 3). This is a clear and comprehensive mission, as all of the graduates are being prepared for the same career, yet it functions to focus the goals and activities throughout the institution. This mission guides the student learning outcomes of six developmental domains: intellectual, military, physical, social, moral/ethical, and the domain of the human spirit (USMA, 2009). The description of activities supporting the achievement of these outcomes is based on an underlying theoretical model which describes identity development as progressing from self-interest through reciprocal exchange, belonging, and on to a personal code of conduct (Kegan, 1982; Keith, 2010).

The intellectual domain is the major focus of the academic program, and is further defined by ten general education goals, described in Figure 3.1. USMA has been engaged in a robust process of assessing these general education goals of the academic program since 1989, the year in which teams of faculty leaders under the support and guidance of the academic dean defined the academic program goals and corresponding outcome statements. In 1996 a series of indirect measures in the form of surveys were developed with items that aligned with the outcome statements from the ten goals, asking freshmen and seniors their level of confidence in performing the skills defined in the outcome statements. In 1997 these questions were also posed during focus group interviews with senior military officers who were in graduate school, and in 1999 the items were incorporated into a survey for graduates and their supervisors at a point three years after graduation.

<table>
<thead>
<tr>
<th>USMA Mission Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>To educate, train, and inspire the Corps of Cadets so that each graduate is a commissioned leader of character committed to the values of Duty, Honor, Country and prepared for a career of professional excellence and service to the Nation as an officer in the United States Army.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overarching Academic Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graduates anticipate and respond effectively to the uncertainties of a changing technological, social, political, and economic world.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Academic Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graduates of the Academy will:</td>
</tr>
<tr>
<td>- think and act creatively.</td>
</tr>
<tr>
<td>- recognize moral issues and apply ethical considerations in decision making.</td>
</tr>
<tr>
<td>- demonstrate the capability and desire to pursue progressive and continued intellectual development.</td>
</tr>
<tr>
<td>- listen, read, speak, and write effectively.</td>
</tr>
<tr>
<td>- draw from an appreciation of culture to understand in a global context human behavior, achievement, and ideas.</td>
</tr>
<tr>
<td>- draw on an appreciation of history to understand in a global context human behavior, achievement and ideas.</td>
</tr>
<tr>
<td>- understand patterns of human behavior, particularly how individuals, organizations, and societies pursue social, political, and economic goals.</td>
</tr>
<tr>
<td>- be scientifically literate and capable of applying scientific, mathematical, and computational modes of thought to the solution of complex problems.</td>
</tr>
<tr>
<td>- apply mathematics, science, technology, and the engineering design process to devise technological problem solutions that are effective and adaptable.</td>
</tr>
<tr>
<td>- understand and apply Information Technology concepts to acquire, manage, communicate and defend information, solve problems, and adapt to technological change.</td>
</tr>
</tbody>
</table>

*Figure 3.1 General education goals*
Since its first publication in 1997, the student learning outcomes have been codified in a document, *Educating Future Army Officers for a Changing World* (USMA, 2007). These are general education goals that are not course specific, but are infused across the set of 30 core courses and reinforced in the majors. Figure 3.1 presents a summary of the USMA Mission, the overarching academic goal, and the ten academic student learning outcome goals that support both the Mission and the overarching academic goal. Figure 3.2 gives an example of the Cultural Perspective goal statement and the student learning outcome statements. These statements were developed by faculty teams to clearly define the expected outcomes of the curriculum. An example of a character goal, also in Figure 3.2, lists the outcomes for the Moral/Ethical Goal.

In 2002 the faculty teams developed student learning outcome standards for each of the outcome statements to use for assessment. Figure 3.2 also notes the results of an analysis of a course content and student learning outcomes matrix which was prepared by the teams. Thirteen of the 30 core courses have activities directly related to one or more of the student learning outcomes of the Cultural Perspective Goal, and three courses have activities supporting the Moral/Ethical Goal. Figure 3.3 provides a more complete description of the standards of the student learning outcomes for the Cultural Perspective Goal, and the courses that were identified with activities contributing to the development of the student learning outcomes.

<table>
<thead>
<tr>
<th>Examples of goals and outcomes at USMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultural Perspective</td>
</tr>
<tr>
<td>Goal Statement</td>
</tr>
<tr>
<td>Outcome statements</td>
</tr>
<tr>
<td>Curriculum alignment mapping</td>
</tr>
<tr>
<td>Direct measures</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Examples of goals and outcomes at USMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultural Perspective</td>
</tr>
<tr>
<td>Indirect measures</td>
</tr>
<tr>
<td>Interview/focus groups</td>
</tr>
</tbody>
</table>

*Figure 3.2 Goals and outcomes*
WHAT GRADUATES CAN DO

DEMONSTRATE ABILITY TO:

WHAT CAN STUDENTS DO

Analyze contemporary and historical events from different cultural perspectives.

Analyze and understand the current and past events of a culture under study as well as that culture’s distinctive interpretation(s) of those events. Complete this analysis by examining traditional responses to change, delineating the components of that culture, regional variations in cultural diversity/human behavior, and spatial linkages among culture systems.

Analyze and understand cultural landscapes (i.e., physical, economic, military, social, cultural and religious systems), their geographic distribution(s), and be cognizant of similarities and differences as they affect decision-making processes.

Analyze and delineate cultural diversity and human behavior in domestic and international scenarios, recognize a range of similarities and differences in common cultural practices, illustrate regional dissimilarity in cultural practices and demonstrate how these differences/similarities drive choices in domestic and international relations and military/civilian law.

View the world from the perspective of someone in another culture.

Analyze and understand the current and past events of the culture under study, and be able to view events, issues, and choices in international relations from multiple perspectives. Understand how individuals and cultures pursue social, political, and economic goals.

Student learning outcome assessment began in earnest as the teams systematically collected the evidence that had been identified as embedded indicators in the course products. The direct measures in Figure 3.3 show that after constructing the course and student learning matrices, faculty identified several course products that could reasonable serve as sources for collecting direct assessment evidence. Note that for the Moral/Ethical Goal, there is a combination of embedded indicators: final exams, papers, and tests- and the add-on indicators of the Defining Issues Test (DIT), and the ALERT (a locally developed version of the DIT with military scenarios).

Student learning outcome assessment began in earnest as the teams systematically collected the evidence that had been identified as embedded indicators in the course products. The last section of Figure 3.2 describes the sources of indirect indicators: surveys of students, graduates and their supervisors, and interviews, with examples of the survey items that are aligned with the student learning outcomes. The students, graduates and supervisors are asked to indicate their level of confidence in these abilities. Figure 3.4 is a rubric developed by the goal team that was applied to a direct embedded indicator: a paper prepared by a student for a core course in physical geography.

When the teams combine the evidence from the embedded indicators with the survey and interview data, they are able to see how well students and graduates are achieving the student learning outcomes for each of the goals. This combination of indirect indicators in the form of surveys and interviews, and direct evidence from student products, created a comprehensive data set for the triangulation of evidence to assess the student learning outcomes (Judd & Keith, 2009; Keith et al., 2002).

To bring the assessment cycle full circle, each goal team prepares a year-end report describing the assessment data collection activities, analyzing the results, identifying concerns, and recommending any changes deemed appropriate to the curriculum or the assessment process. The reports from the goal teams are summarized for the academic area and forwarded to the chief academic officer for review and action.

Over the course of several years, the Cultural Perspective goal team came to the consistent conclusion that while students are achieving the learning outcomes at a basic level, student achievement of the outcomes could be improved if the curriculum included the development of a framework for

In 2003 the teams identified course products throughout the core curriculum that could be used as embedded indicators for each of the student learning outcome statements. The direct measures in Figure 3.3 show that after constructing the course and student learning matrices, faculty identified several course products that could reasonable serve as sources for collecting direct assessment evidence. Note that for the Moral/Ethical Goal, there is a combination of embedded indicators: final exams, papers, and tests- and the add-on indicators of the Defining Issues Test (DIT), and the ALERT (a locally developed version of the DIT with military scenarios).

Student learning outcome assessment began in earnest as the teams systematically collected the evidence that had been identified as embedded indicators in the course products. The direct measures in Figure 3.3 show that after constructing the course and student learning matrices, faculty identified several course products that could reasonable serve as sources for collecting direct assessment evidence. Note that for the Moral/Ethical Goal, there is a combination of embedded indicators: final exams, papers, and tests- and the add-on indicators of the Defining Issues Test (DIT), and the ALERT (a locally developed version of the DIT with military scenarios).

Figure 3.3 Cultural perspective outcomes and embedded indicators
understanding a new cultural environment. These findings inspired discussions across the institution, and an examination of areas inside and outside the academic curriculum where the development of cultural understanding could be reinforced. In fact, numerous opportunities were identified in the academic curriculum (e.g., in language and history course coordination; psychology course treatment of cultural differences) and the military training that occurs during the summer months (Keith, 2010).

In response to the recommendations made by the assessment team, the chief academic officer asked the assessment group to define what an interculturally competent student learning outcome should be, what is required to achieve that outcome, what elements of the current curriculum support the development of intercultural competence, and the feasibility of providing every student with an intercultural experience prior to graduation. In the true spirit of the assessment cycle, this is still a work in progress, as the institution assesses, recommends, responds, and assesses again.

Figure 3.4 Cultural perspective rubric for assessing embedded indicators
Lessons Learned From the Assessment of Student Learning Outcomes

As articulated by Keith (2010), we have learned that the establishment of the student learning outcomes has defined an educated graduate from USMA, and described the learning system we have implemented to achieve those outcomes.

We have learned the value of involving all of the stakeholders in order to have universal acceptance of the system. The first iteration showed that without broad involvement across the institution in developing the model there was little sense of awareness or ownership among faculty and staff.

We have learned that the existing compartmentalization of programs has to be breached if students are expected to integrate learning across the curriculum and programs. Students tend to be risk averse, afraid of failure, and unwilling to leave their comfort zones. Therefore, we cannot wish for integration to take place; we must intentionally set the conditions and expectations for integration, and remove the institutional barriers.

We have learned that there may be limitations in how far students can be expected to develop along the model we have embraced (Lewis et al., 2005). Nevertheless, for students to continue to develop to the maximum extent while they are with us, they need to have the opportunities to assess their own performance structured into the learning system.

Our chosen learning model views development as continuing well after graduation, hence the title of our learning system begins with the words “building capacity,” in the recognition that we are setting the stage for students’ development well into the early years of their careers.

The process of developing and implementing a cyclical system of assessment of student learning outcomes has helped USMA determine what we intend to accomplish with our students, how we intend to accomplish it, how well we are accomplishing it, and how we can improve the preparation of our students for their careers.

References


Banta, T. W., & Pike, G. R. (this volume). The bottom line: Will faculty USE assessment findings?


Pitoniak, M. J., & Morgan, D. L. (this volume). Setting and validating cut scores for tests.