

### Rubric for Evaluating Institutional Student Learning Assessment Processes

This is intended for institutions to use as a tool to help them assess the status of their **current assessment efforts in terms of Middle States' accreditation standards and expectations**. This tool is **not** intended to be used by any evaluators or to prescribe specific Commission actions regarding the institution.

- No plans = **No documented evidence that the institution has plans** to do this.
- No evidence = The institution appears to be aware that it should do this, but there is **no documented evidence that this is happening**.
- A few areas = The institution has documented evidence that this is happening **in just a few areas** (for example, only in programs with specialized accreditation).
- Some areas = The institution has documented evidence—not just assurances—that this is happening **in some but not most areas** (for example, in a number of academic programs but not yet in general education)
- Most areas = The institution has documented evidence—not just assurances—that this is happening **in most but not all areas**.
- Everywhere = The institution has documented evidence—not just assurances—that this is happening **everywhere**.

	<i>For academic programs, the general education curriculum, and institutional goals articulated in the mission statement, vision statement, or elsewhere:</i>	No plans	No evidence	A few areas	Some areas	Most areas	Everywhere
1	Institutional leaders demonstrate sustained—not just one-time or periodic—support for promoting an ongoing culture of assessment and for efforts to improve teaching.						
2	Clear statements of expected learning outcomes at the institutional, unit, program, and course levels have been developed and have appropriate interrelationships.						
3	Those with a vested interest in the learning outcomes of the institution, program, or curriculum are involved in developing, articulating, and assessing them.						
4	Statements of program-level expected learning outcomes are made available to current and prospective students.						
5	Course syllabi include statements of expected learning outcomes.						
6	Targets or benchmarks for determining whether student learning outcomes have been achieved have been established and justified; the justifications demonstrate that the targets are of appropriate college-level rigor <b>and are appropriate given the institution's mission</b> .						
7	Multiple measures of student learning, including direct evidence, have been collected and are of sufficient quality that they can be used with confidence to make appropriate decisions.						
8	The evidence of student learning that has been collected is clearly linked to expected learning outcomes.						
9	Student learning assessment results have been shared in useful forms and discussed with appropriate constituents, including those who can effect change.						
10	Student learning assessment results have been used to improve teaching and by institutional leaders to inform planning and budgeting decisions.						
11	In any areas in which the above are not yet happening, concrete, feasible, and timely plans are in place.						
12	Assessment processes have been reviewed and changes have been made to improve their effectiveness and/or efficiency, as appropriate.						
13	There is sufficient engagement, momentum, and simplicity in current assessment practices to provide assurance that assessment processes will be sustained indefinitely.						

## From Teaching to Learning - A New Paradigm for Undergraduate Education

By Robert B. Barr and John Tagg

The article we are including originally appeared in the November/December 1995 edition of *Change* magazine. In the short time it has been out, it has created considerable discussion on educational list serves and has prompted numerous requests for reprints in newsletters.

*(Robert B Barr is director of institutional research and planning and John Tagg is associate professor of English at Palomar College, San Marcos California.)*

---

*The significant problems we face cannot be solved at the same level of thinking we were at when we created them. -Albert Einstein*

A paradigm shift is taking hold in American higher education. In its briefest form, the paradigm that has governed our colleges is this: A college is an institution that exists *to provide* instruction. Subtly but profoundly we are shifting to a new paradigm: A college is an institution that exists *to produce learning*. This shift changes everything. It is both needed and wanted. *(See chart comparing two paradigms)*

We call the traditional, dominant paradigm the "Instruction Paradigm." Under it, colleges have created complex structures to provide for the activity of teaching conceived primarily as delivering 50-minute lectures-the mission of a college is to deliver instruction.

Now, however, we are beginning to recognize that our dominant paradigm mistakes a means for an end. It takes the means or method-called "instruction" or "teaching"-and makes it the college's end or purpose. To say that the purpose of colleges is to provide instruction is like saying that General Motors' business is to operate assembly lines or that the purpose of medical care is to fill hospital beds. We now see that our mission is not instruction but rather that of producing *learning* with every student by *whatever* means work best.

The shift to a "Learning Paradigm" liberates institutions from a set of difficult constraints. Today it is virtually impossible for them to respond effectively to the challenge of stable or declining budgets while meeting the increasing demand for post secondary education from increasingly diverse students. Under the logic of the Instruction Paradigm, colleges suffer from a serious design flaw: it is not possible to increase outputs without a corresponding increase in costs, because any attempt to increase outputs without increasing resources is a threat to quality. If a college attempts to increase its productivity by increasing either class sizes or faculty workloads, for example, academics will be quick to assume inexorable negative consequences for educational quality.

Just as importantly, the Instruction Paradigm rests on conceptions of teaching that are

increasingly recognized as ineffective. As Alan Guskin pointed out in a September/October 1994 *Change* article premised on the shift from teaching to learning, "the primary learning environment for undergraduate students, the fairly passive lecture-discussion format where faculty talk and most students listen, is contrary to almost every principle of optimal settings for student learning." The Learning Paradigm ends the lecture's privileged position, honoring in its place whatever approaches serve best to prompt learning of particular knowledge by particular students.

The Learning Paradigm also opens up the truly inspiring goal that each graduating class learns more than the previous graduating class. In other words, the Learning Paradigm envisions the institution itself as a learner- over time, it continuously learns how to produce more learning with each graduating class, each entering student.

For many of us, the Learning Paradigm has always lived in our hearts. As teachers, we want above all else for our students to learn and succeed. But the heart's feeling has not lived clearly and powerfully in our heads. Now, as the elements of the Learning Paradigm permeate the air. Our heads are beginning to understand what our hearts have known. However, none of us has yet put all the elements of the Learning Paradigm together in a conscious, integrated whole.

Lacking such a vision, we've witnessed reformers advocate many of the new paradigm's elements over the years, only to see few of them widely adopted. The reason is that they have been applied piecemeal within the structures of a dominant paradigm that rejects or distorts them. Indeed, for two decades the response to calls for reform from national commissions and task forces generally has been an attempt to address the issues *within the* framework of the Instruction Paradigm. The movements thus generated have most often failed, undone by the contradictions within the traditional paradigm. For example, if students are not learning to solve problems or think critically, the old logic says we must teach a class in thinking and make it a general education requirement. The logic is all too circular: What students are learning in the classroom doesn't address their needs or ours; therefore, we must bring them back into another classroom and instruct them some more. The result is never what we hope for because, as Richard Paul, director of the Center for Critical Thinking observes glumly, "critical thinking is taught in the same way that other courses have traditionally been taught, with an excess of lecture and insufficient time for practice."

To see what the Instruction Paradigm is we need only look at the structures and behaviors of our colleges and infer the governing principles and beliefs they reflect. But it is much more difficult to see the Learning Paradigm, which has yet to find complete expression in the structures and processes of any college. So we must imagine it. This is what we propose to do here. As we outline its principles and elements, we'll suggest some of their implications for colleges-but only some, because the expression of principles in concrete structures depends on circumstances. It will take decades to work out many of the Learning Paradigm's implications. But we hope here that by making it more explicit we will help colleagues to more fully recognize it and restructure our institutions in its image.

That such a restructuring is needed is beyond question: the gap between what we *say* we want of higher education and what its structures *provide* has never been wider. To use a distinction made by Chris Argyris and Donald Schon, the difference between our espoused theory and our theory-in-use is becoming distressingly noticeable. An "espoused theory," readers will recall, is the set of principles people offer to explain their behavior; the principles we can infer from how

people or their organizations actually behave is their "theory-in-use." Right now, the Instruction Paradigm is our theory-in-use, yet the *espoused* theories of most educators more closely resemble components of the Learning Paradigm. The more we discover about how the mind works and how students learn, the greater the disparity between what we say and what we do. Thus so many of us feel increasingly constrained by a system increasingly at variance with what we believe. To build the colleges we need for the 21st century-to put our minds where our hearts are, and rejoin acts with beliefs-we must consciously reject the Instruction Paradigm and restructure what we do on the basis of the Learning Paradigm.

(See chart comparing two paradigms)

## THE PARADIGMS

When comparing alternative paradigms, we must take care: the two will seldom be as neatly parallel as our summary chart suggests. A paradigm is like the rules of a game: one of the functions of the rules is to define the playing field and domain of possibilities on that field. But a new paradigm may specify a game played on a larger or smaller field with a larger or smaller domain of legitimate possibilities. Indeed, the Learning Paradigm expands the playing held and domain of possibilities and it radically changes various aspects of the game. In the Instruction Paradigm, a specific methodology determines the boundary of what colleges can do; in the Learning Paradigm, student learning and success set the boundary. By the same token, not all elements of the new paradigm are contrary to corresponding elements of the old; the new includes many elements of the old within its larger domain of possibilities. The Learning Paradigm does not prohibit lecturing, for example. Lecturing becomes one of many possible methods, all evaluated on the basis of their ability to promote appropriate learning.

In describing the shift from an Instruction to a Learning Paradigm, we limit our address in this article to undergraduate education. Research and public service are important functions of colleges and universities but lie outside the scope of the present discussion. Here, as in our summary chart, we'll compare the two paradigms along six dimensions: mission and purposes, criteria for success, teaching/learning structures, learning theory, productivity and funding, and nature of roles.

## MISSION AND PURPOSES

In the Instruction Paradigm, the mission of the college is to provide instruction, to teach. The method and the product are one and the same. The means is the end. In the Learning Paradigm, the mission of the college is to produce learning. The method and the product are separate. The end governs the means.

Some educators may be uncomfortable with the verb "produce." We use it because it so strongly connotes that the college takes *responsibility* for learning. The point of saying that colleges are to *produce* learning-not provide, not support, not encourage-is to say, unmistakably, that they are responsible for the degree to which students learn. The Learning Paradigm shifts what the institution takes responsibility for: from quality instruction (lecturing, talking) to student learning. Students, the co-producers of learning, can and must, of course, take responsibility for their own learning. Hence, responsibility is a win-win game wherein two agents take responsibility for the same outcome even though neither is in complete control of all

the variables. When two agents take such responsibility, the resulting synergy produces powerful results.

The idea that colleges cannot be responsible for learning flows from a disempowering notion of responsibility. If we conceive of responsibility as a fixed quantity in a zero-sum game, then students must take responsibility for their own learning, and no one else can. This model generates a concept of responsibility capable of assigning blame but not of empowering the most productive action. The concept of responsibility as a framework for action is quite different: when one takes responsibility, one sets goals and then acts to achieve them, continuously modifying one's behavior to better achieve the goals. To take responsibility for achieving an outcome is not to guarantee the outcome, nor does it entail the complete control of all relevant variables; it is to make the achievement of the outcome the criterion by which one measures one's own efforts. In this sense, it is no contradiction to say that students, faculty, and the college as an institution can all take responsibility for student learning.

In the Learning Paradigm, colleges take responsibility for learning at two distinct levels. At the organizational level, a college takes responsibility for the aggregate of student learning and success. Did, for example, the graduating class's mastery of certain skills or knowledge meet our high, public standards for the award of the degree? Did the class's knowledge and skills improve over those of prior classes? The college also takes responsibility at the individual level, that is, for each individual student's learning. Did Mary Smith learn the chemistry we deem appropriate for a degree in that field? Thus, the institution takes responsibility for both its institutional outcomes and individual student outcomes.

Turning now to more specific purposes, in the Instruction Paradigm, a college aims to transfer or deliver knowledge from faculty to students; it offers courses and degree programs and seeks to maintain a high quality of instruction within them, mostly by assuring that faculty stay current in their fields. If new knowledge or clients appear, so will new course work. The very purpose of the Instruction Paradigm is to offer courses. In the Learning Paradigm, on the other hand, a college's purpose is not to transfer knowledge but to create environments and experiences that bring students to discover and construct knowledge for themselves, to make students members of communities of learners that make discoveries and solve problems. The college aims, in fact, to create a series of ever more powerful learning environments. The Learning Paradigm does not limit institutions to a single means for empowering students to learn; within its framework, effective learning technologies are continually identified, developed, tested, implemented, and assessed against one another. The aim in the Learning Paradigm is not so much to improve the quality of instruction- although that is not irrelevant-as it is to improve continuously the quality of learning for students individually and in the aggregate.

Under the older paradigm, colleges aimed to provide access to higher education, especially for historically under-represented groups such as African-Americans and Hispanics. Too often, mere access hasn't served students well. Under the Learning Paradigm, the goal for under represented students (and *all* students) becomes not simply access but success. By "success" we mean the achievement of overall student educational objectives such as earning a degree, persisting in school, and learning the "right" things-the skills and knowledge that will help students to achieve their goals in work and life. A Learning Paradigm college, therefore, aims for ever-higher graduation rates while maintaining or even increasing learning standards

By shifting the intended institutional outcome from teaching to learning, the Learning Paradigm makes possible a continuous improvement in productivity. Whereas under the Instruction Paradigm a primary institutional purpose was to optimize faculty well-being and success—including recognition for research and scholarship—in the Learning Paradigm a primary drive is to produce learning outcomes more efficiently. The philosophy of an Instruction Paradigm college reflects the belief that it cannot increase learning outputs without more resources, but a Learning Paradigm college expects to do so continuously. A Learning Paradigm college is concerned with learning productivity, not teaching productivity .

## CRITERIA FOR SUCCESS

Under the Instruction Paradigm, we judge our colleges by comparing them to one another. The criteria for quality are defined in terms of inputs and process measures. Factors such as selectivity in student admissions, number of PhDs on the faculty, and research reputation are used to rate colleges and universities. Administrators and boards may look to enrollment and revenue growth and the expansion of courses and programs. As Guskin put it, "We are so wedded to a definition of quality based on resources that we find it extremely difficult to deal with the *results* of our work, namely student learning."

The Learning Paradigm necessarily incorporates the perspectives of the assessment movement. While this movement has been under way for at least a decade, under the dominant Instruction Paradigm it has not penetrated very far into normal organizational practice. Only a few colleges across the country systematically assess student learning outcomes. Educators in California community colleges always seem to be surprised when they hear that 45 percent of first-time fall students do not return in the spring and that it takes an average of six years for a student to earn an associate's (AA) degree. The reason for this lack of outcomes knowledge is profoundly simple: under the Instruction Paradigm, student outcomes are simply irrelevant to the successful functioning and funding of a college.

Our faculty evaluation systems, for example, evaluate the performance of faculty in teaching terms, not learning terms. An instructor is typically evaluated by her peers or dean on the basis of whether her lectures are organized, whether she covers the appropriate material, whether she shows interest in and understanding of her subject matter, whether she is prepared for class, and whether she respects her students' questions and comments. All these factors evaluate the instructor's performance in teaching terms. They do not raise the issue of whether students are learning, let alone demand evidence of learning or provide for its reward.

Many institutions construe teaching almost entirely in terms of lecturing. A true story makes the point. A biology instructor was experimenting with collaborative methods of instruction in his beginning biology classes. One day his dean came for a site visit, slipping into the back of the room. The room was a hubbub of activity. Students were discussing material enthusiastically in small groups spread out across the room; the instructor would observe each group for a few minutes, sometimes making a comment, sometimes just nodding approval. After 15 minutes or so the dean approached the instructor and said, "I came today to do your evaluation. I'll come back another time when you're teaching."

In the Instruction Paradigm, teaching is judged on its own terms; in the Learning Paradigm, the power of an environment or approach is judged in terms of its impact on learning. If learning

occurs, then the environment has power. If students learn more in environment A than in environment B, then A is more powerful than B. To know this in the Learning Paradigm we would assess student learning routinely and constantly.

Institutional outcomes assessment is analogous to classroom assessment, as described by K. Patricia Cross and Thomas Angelo. In our own experience of classroom-assessment training workshops, teachers share moving stories about how even limited use of these techniques has prompted them to make big changes in their teaching, sometimes despite years of investment in a previous practice. Mimi Steadman, in a recent study of community college teachers using classroom assessment, found that "eighty-eight percent of faculty surveyed reported that they had made changes in their teaching behaviors as a result." This at first was startling to us. How could such small amounts of information produce such large changes in teacher behavior? Upon reflection, it became clear. The information was feedback about learning, about results—something teachers rarely collect. Given information that their students were not learning, it was obvious to these teachers that something had to be done about the methods they had been using. Likewise, we think, feedback on learning results at the institutional level should have a correspondingly large impact on an institution's behavior and on the means it uses to produce learning.

Of course, some will argue, true education simply cannot be measured. You cannot measure, for example, true appreciation of the beauty of a work of art. Certainly some learning is difficult, even impossible to measure. But it does not follow that useful and meaningful assessment is impossible.

If we compare outcomes assessment with the input measures controlling policy in the Instruction Paradigm, we find that measures of outcome provide far more genuine information about learning than do measures of input. Learning outcomes include whatever students do as a result of a learning experience. Any measurement of students' products from an educational experience is a measure of a learning outcome. We could count the number of pages students write, the number of books they read, their number of hours at the computer, or the number of math problems they solve.

Of course, these would be silly methods to determine institutional incentives, and we do not recommend them. Any one of them, however, would produce more useful information on learning than the present method of measuring inputs and ignoring outcomes. It would make more sense to fund a college on the number of math problems students solve, for example, than to fund it on the number of students who sit in math classes. We suspect that *any* system of institutional incentives based on outcomes would lead to greater learning than any system of incentives based on inputs. But we need not settle for a system biased toward the trivial. Right now, today, we can construct a good assessment regime with the tools we have at hand.

The Learning Paradigm requires us to heed the advice of the Wingspread Group: "New forms of assessment should focus on establishing what college and university graduates have learned—the knowledge and skill levels they have achieved and their potential for further independent learning."

## **TEACHING/LEARNING STRUCTURES**

By structures we mean those features of an organization that are stable over time and that form the framework within which activities and processes occur and through which the purposes of the organization are achieved. Structure includes the organization chart, role and reward systems, technologies and methods, facilities and equipment, decision-making customs, communication channels, feedback loops, financial arrangements, and funding streams.

Peter Senge, in *The Fifth Discipline*, a book about applying systems theory to organizational learning, observes that institutions and their leaders rarely focus their attention on systemic structures. They seldom think, he says, to alter basic structures in order to improve organizational performance, even though those structures generate the patterns of organizational action and determine which activities and results are possible. Perhaps the recent talk about restructuring, reengineering, and reinvention in higher education reflects a change in focus and a heightened awareness of both the constraining and liberating power of organizational structures.

There is good reason to attend to structure. First, restructuring offers the greatest hope for increasing organizational efficiency and effectiveness. Structure is leverage. If you change the structure in which people work, you increase or decrease the leverage applied to their efforts. A change in structure can either increase productivity or change the nature of organizational outcomes. Second, structure is the concrete manifestation of the abstract principles of the organization's governing paradigm. Structures reflecting an old paradigm can frustrate the best ideas and innovations of new-paradigm thinkers. As the governing paradigm changes, so likewise must the organization's structures.

In this section, we focus on the main structures related to the teaching and learning process; funding and faculty role structures are discussed later under separate headings.

The teaching and learning structure of the Instruction Paradigm college is atomistic. In its universe, the "atom" is the 50-minute lecture, and the "molecule" is the one-teacher, one-classroom, three-credit hour course. From these basic units the physical architecture, the administrative structure, and the daily schedules of faculty and students are built. Dennis McGrath and Martin Spear, professors at the Community College of Philadelphia, note that "education proceeds everywhere through the vehicle of the three-credit course. Faculty members land everyone else, we might add I have so internalized that constraint that they are long past noticing that it is a constraint, thinking it part of the natural order of things."

The resulting structure is powerful and rigid. It is, of course, perfectly suited to the Instruction Paradigm task of offering one-teacher, one-classroom courses. It is antithetical to creating almost any other kind of learning experience. A sense of this can be obtained by observing the effort, struggle, and rule bending required to schedule even a slightly different kind of learning activity, such as a team-taught course.

In the "educational atomism" of the Instruction Paradigm, the parts of the teaching and learning process are seen as discrete entities. The parts exist prior to and independent of any whole; the whole is no more than the sum of the parts, or even less. The college interacts with students only in discrete, isolated environments, cut off from one another because the parts—the classes—are prior to the whole. A "college education" is the sum the student's experience of a series of discrete, largely unrelated, three-credit classes.



In the Instruction Paradigm, the teaching and learning process is governed by the further rule that time will be held constant while learning varies. Although addressing public elementary and secondary education, the analysis of the National Commission on Time and Learning nonetheless applies to colleges:

*Time is learning's warden. Our time-bound mentality has fooled us all into believing that schools can educate all of the people all of the time in a school year of 180 six-hour days...If experience, research, and common sense teach nothing else, they confirm the truism that people learn at different rates, and in different ways with different subjects. But we have put the cart before the horse: our schools...are captives of clock and calendar. The boundaries of student growth are defined by schedules... instead of standards for students and learning.*

Under the rule of time, all classes start and stop at the same time and take the same number of calendar weeks. The rule of time and the priority of parts affect every instructional act of the college.

Thus it is, for example, that if students come into college classes "unprepared," it is not the job of the faculty who teach those classes to "prepare" them. Indeed, the structure of the one-semester, three-credit class makes it all but impossible to do so. The only solution, then, is to create new courses to prepare students for the existing courses; within the Instruction Paradigm, the response to educational problems is always to generate more atomized, discrete instructional units. If business students are lacking a sense of ethics, then offer and require a course in business ethics. If students have poor study skills, then offer a "master student" course to teach such skills.

Instruction Paradigm colleges atomistically organize courses and teachers into departments and programs that rarely communicate with one another. Academic departments, originally associated with coherent disciplines, are the structural home bases for accomplishing the essential work of the college: offering courses. "Departments have a life of their own," notes William D. Schaefer, professor of English and former executive vice chancellor at UCLA. They are "insular, defensive, self-governing, [and] compelled to protect their interests because the faculty positions as well as the courses that justify funding those positions are located therein."

Those globally applicable skills that are the foundation of meaningful engagement with the world—reading, writing, calculating, reasoning—find a true place in this structure only if they have their own independent bases: the English or math or reading departments. If students cannot reason or think well, the college creates a course on reasoning and thinking. This in turn produces pressure to create a corresponding department. "If we are not careful," warns Adam Sweeting, director of the Writing Program at the Massachusetts School of Law at Andover, "the teaching of critical thinking skills will become the responsibility of one university department, a prospect that is at odds with the very idea of a university."

Efforts to extend college-level reading, writing, and reasoning "across the curriculum" have largely failed. The good intentions produced few results because, under the Instruction Paradigm, the teacher's job is to "cover the material" as outlined in the disciplinary syllabus. The instructor charged with implementing writing or reading or critical thinking "across the curriculum" often must choose between doing her job or doing what will help students learn—between doing well, as it were, or doing good.

From the point of view of the Learning Paradigm, these Instruction Paradigm teaching and learning structures present immense barriers to improving student learning and success. They provide no space and support for redesigned learning environments or for experimenting with alternative learning technologies. They don't provide for, warrant, or reward assessing whether student learning has occurred or is improving.

In a Learning Paradigm college, the structure of courses and lectures becomes dispensable and negotiable. Semesters and quarters, lectures, labs, syllabi-indeed, classes themselves- become options rather than received structures or mandatory activities. The Learning Paradigm prescribes no one "answer" to the question of how to organize learning environments and experiences. It supports any learning method and structure that works, where "works" is defined in terms of learning outcomes, not as the degree of conformity to an ideal classroom archetype. In fact, the Learning Paradigm requires a constant search for new structures and methods that work better for student learning and success, and expects even these to be redesigned continually and to evolve over time.

The transition from Instruction Paradigm to Learning Paradigm will not be instantaneous. It will be a process of gradual modification and experimentation through which we alter many organizational parts in light of a new vision for the whole. Under the Instruction Paradigm, structures are assumed to be fixed and immutable; there is no ready means for achieving the leverage needed to alter them. The first structural task of the Learning Paradigm, then, is to establish such leverage.

The key structure for changing the rest of the system is an institution wide assessment and information system- an essential structure in the Learning Paradigm, and a key means for getting there. It would provide constant, useful feedback on institutional performance. It would track transfer, graduation, and other completion rates. It would track the flow of students through learning stages (such as the achievement of basic skills) and the development of in-depth knowledge in a discipline. It would measure the knowledge and skills of program completers and graduates. It would assess learning along many dimensions and in many places and stages in each student's college experience.

To be most effective, this assessment system would provide public institutional-level information. We are not talking about making public the status of individual students by name, but about making the year-to-year graduation rate-or the mean score of graduating seniors on a critical thinking assessment, for example-"public" in the sense that they are available to everyone in the college community. Moreover, in the Learning Paradigm college, such data are routinely talked about and acted upon by a community ever dedicated to improving its own performance.

The effectiveness of the assessment system for developing alternative learning environments depends in part upon its being *external* to learning programs and structures. While in the Instruction Paradigm students are assessed and graded within a class by the same instructor responsible for teaching them, in the Learning Paradigm much of the assessment would be independent of the learning experience and its designer, somewhat as football games are independent measures of what is learned in football practice. Course grades alone fail to tell us what students know and can do; average grades assigned by instructors are not reliable measures of whether the institution is improving learning.

Ideally, an institution's assessment program would measure the "value-added" over the course of students' experience at the college. Student knowledge and skills would be measured upon entrance and again upon graduation, and at intermediate stages such as at the beginning and completion of major programs. Students could then be acknowledged and certified for what they have learned; the same data, aggregated, could help shift judgments of institutional quality from inputs and resources to the value-added brought to student learning by the college.

The college devoted to learning first identifies the knowledge and skills it expects its graduates to possess, without regard to any particular curriculum or educational experiences. It then determines how to assess them reliably. It assesses graduating students, and the resulting information is then used to redesign and improve the processes and environments leading to such outcomes. In this manner, enhancing intellectual skills such as writing and problem solving and social skills such as effective team participation become the project of *all* learning programs and structured experiences. The whole would govern the parts.

Information from a sophisticated assessment system will gradually lead to the transformation of the college's learning environments and supporting structures. Such a system seeks out "best practice" benchmarks against which improvements in institutional performance can be measured in learning terms. It is the foundation for creating an institutional capacity to develop ever more effective and efficient ways of empowering learning. It becomes the basis for generating revenue or funding according to learning results rather than hours of instruction. Most importantly, it is the key to the college's and its staff's taking responsibility for and enjoying the progress of each student's education.

Instead of fixing the means-such as lectures and courses-the Learning Paradigm fixes the ends, the learning results, allowing the means to vary in its constant search for the most effective and efficient paths to student learning. Learning outcomes and standards thus would be identified and held to for all students-or *raised* as learning environments became more powerful- while the time students took to achieve those standards would vary. This would reward skilled and advanced students with speedy progress while enabling less prepared students the time they needed to actually master the material. By "testing out," students could also avoid wasting their time being "taught" what they already know. Students would be given "credit" for degree-relevant knowledge and skills regardless of how or where or when they learned them.

In the Learning Paradigm, then, a college degree would represent not time spent and credit hours dutifully accumulated, but would certify that the student had demonstrably attained specified knowledge and skills. Learning Paradigm institutions would develop and publish explicit exit standards for graduates and grant degrees and certificates only to students who met them. Thus colleges would move away from educational atomism and move toward treating holistically the knowledge and skills required for a degree.

## **LEARNING THEORY**

The Instruction Paradigm frames learning atomistically. In it, knowledge, by definition, consists of matter dispensed or delivered by an instructor. The chief agent in the process is the teacher who delivers knowledge; students are viewed as passive vessels, ingesting knowledge for recall on tests. Hence, any expert can teach. Partly because the teacher knows which chunks of

knowledge are most important, the teacher controls the learning activities. Learning is presumed to be cumulative because it amounts to ingesting more and more chunks. A degree is awarded when a student has received a specified amount of instruction.

The Learning Paradigm frames learning holistically, recognizing that the chief agent in the process is the learner. Thus, students must be active discoverers and constructors of their own knowledge. In the Learning Paradigm, knowledge consists of frameworks or wholes that are created or constructed by the learner. Knowledge is not seen as cumulative and linear, like a wall of bricks, but as a nesting and interacting of frameworks. Learning is revealed when those frameworks are used to understand and act. Seeing the whole of something-the forest rather than the trees, the image of the newspaper photo rather than its dots- gives meaning to its elements, and that whole becomes more than a sum of component parts. Wholes and frameworks can come in a moment-a flash of insight-often after much hard work with the pieces, as when one suddenly knows how to ride a bicycle.

In the Learning Paradigm, learning environments and activities are learner-centered and learner-controlled. They may even be "teacherless." While teachers will have designed the learning experiences and environments students use-often through teamwork with each other and other staff-they need not be present for or participate in every structured learning activity.

Many students come away from college with a false notion of what learning is and come to believe falsely that learning-at least for some subjects-is too difficult for them. Many students cruise through schools substituting an ersatz role-playing exercise for learning.

The first time I (Barr) studied calculus as a college freshman, I did well by conventional standards. However, while I could solve enough problems to get A's on exams, I really didn't feel that I understood the Limit Theorem, the derivative, or much else. But 15 years later, after having completed college and graduate school and having taught algebra and geometry in high school, I needed to relearn calculus so that I could tutor a friend. In only two, albeit intense, days, I relearned-or really learned for the first time, so it seemed- two semesters of calculus. During those days, I wondered how I ever thought calculus was difficult and why I didn't see the Limit Theorem and derivative for the simple, obvious things they are.

What was the difference between my first learning of calculus and the second? It certainly wasn't a higher IQ. And I don't think it was because I learned or remembered much from the first time. I think it was that I brought some very powerful intellectual frameworks to the learning the second time that I didn't have the first time. Having taught algebra and geometry, I had learned their basic structure, that is, the nature of a mathematical system. I had learned the lay of the land, the whole. Through many years of schooling and study, I had also learned a number of other frameworks that were useful for learning calculus. Thus learning calculus the second time within these "advanced" frameworks was easy compared to learning, or trying to learn, calculus without them as I did as a freshman.

So much of this is because the "learning" that goes on in Instruction Paradigm colleges frequently involves only rudimentary, stimulus-response relationships whose cues may be coded into the context of a particular course but are not rooted in the student's everyday, functioning understanding.

The National Council on Vocational Education summarizes the consequences in its 1991 report, *Solutions*: "The result is fractionation, or splitting into pieces: having to learn disconnected sub-routines, items, and sub-skills without an understanding of the larger context into which they fit and which gives them meaning." While such approaches are entirely consistent with educational atomism, they are at odds with the way we think and learn. The same report quotes Sylvia Farnham-Diggory's summary of contemporary research: "Fractionated instruction maximizes forgetting, inattention, and passivity. Both children and adults acquire knowledge from active participation in holistic, complex, meaningful environments organized around long-term goals. Today's school programs could hardly have been better designed to prevent a child's natural learning system from operating."

The result is that when the contextual cues provided by the class disappear at the end of the semester, so does the learning. Howard Gardner points out that "researchers at Johns Hopkins, MIT, and other well-regarded universities have documented that students who receive honor grades in college-level physics courses are frequently unable to solve basic problems and questions encountered in a form slightly different form that on which they have been formally instructed and tested."

The Learning Paradigm embraces the goal of promoting what Gardner calls "education for understanding"- "a sufficient grasp of concepts, principles, or skills so that one can bring them to bear on new problems and situations, deciding in which ways one's present competencies can suffice and in which ways one may require new skills or knowledge." This involves the mastery of functional, knowledge-based intellectual frameworks rather than the short-term retention of fractionated, contextual cues.

The learning theory of the Instruction Paradigm reflects deeply rooted societal assumptions about talent, relationships, and accomplishment: that which is valuable is scarce; life is a win-lose proposition; and success is an individual achievement. The Learning Paradigm theory of learning reverses these assumptions.

Under the Instruction Paradigm, faculty classify and sort students, in the worst cases into those who are "college material" and those who cannot "cut it," since intelligence and ability are scarce. Under the Learning Paradigm, faculty- and everybody else in the institution-are unambiguously committed to each student's success. The faculty and the institution take an R. Buckminster Fuller view of students: human beings are born geniuses and designed for success. If they fail to display their genius or fail to succeed, it is because their design function is being thwarted. This perspective is founded not in wishful thinking but in the best evidence about the real capabilities of virtually all humans for learning. As the Wingspread Group points out, "There is growing research evidence that all students can learn to much higher standards than we now require." In the Learning Paradigm, faculty find ways to develop every student's vast talents and clear the way for every student's success.

Under the Instruction Paradigm, the classroom is competitive and individualistic, reflecting a view that life is a win-lose proposition. The requirement that the students must achieve individually and solely through their own efforts reflects the belief that success is an individual accomplishment. In the Learning Paradigm, learning environments-while challenging-are win-win environments that are cooperative, collaborative, and supportive.

They are designed on the principle that accomplishment and success are the result of teamwork and group efforts, even when it appears one is working alone.

## **PRODUCTIVITY AND FUNDING**

Under the Instruction Paradigm, colleges suffer from a serious design flaw- they are structured in such a way that they cannot increase their productivity without diminishing the quality of their product. In the Instruction Paradigm, productivity is defined as cost per hour of instruction per student. In this view, the very quality of teaching and learning is threatened by any increase in the student-to-faculty ratio.

Under the Learning Paradigm, productivity is redefined as the cost per unit of learning per student. Not surprisingly, there is as yet no standard statistic that corresponds to this notion of productivity. Under this new definition, however, it is possible to increase outcomes without increasing costs. An abundance of research shows that alternatives to the traditional semester-length, classroom-based lecture method produce more learning. Some of these alternatives are less expensive; many produce more learning for the same cost. Under the Learning Paradigm, producing more with less becomes possible because the more that is being produced is learning and not hours of instruction. Productivity, in this sense, cannot even be measured in the Instruction Paradigm college. All that exists is a measure of exposure to instruction.

Given the Learning Paradigm's definition, increases in productivity pose no threat to the quality of education. Unlike the current definition, this new definition requires that colleges actually produce learning. Otherwise, there is no "product" to count in the productivity ratio.

But what should be the definition of "unit of learning" and how can it be measured? A single, permanent answer to that question does not and need not exist. We have argued above that learning, or at least the effects of learning, can be measured, certainly well enough to determine what students are learning and whether the institution is getting more effective and efficient at producing it.

The Instruction Paradigm wastes not only institutional resources but the time and energy of students. We waste our students' time with registration lines, bookstore lines, lock step class scheduling, and redundant courses and requirements. We do not teach them to learn efficiently and effectively. We can do a lot, as D. Bruce Johnstone, former chancellor of SUNY, suggests, to reduce the false starts and aimless "drift" of students that slow their progress toward a degree.

Now let's consider how colleges are funded. One of the absurdities of current funding formulas is that an institution could utterly fail its educational mission and yet its revenue would remain unaffected. For example, attendance at public colleges on the semester system is measured twice, once in the fall and again in the spring. Normally, at California community colleges, for example, about two-thirds of fall students return for the spring term. New students and returning stop-outs make up for the one-third of fall students who leave. Even if only half-or none at all-returned, as long as spring enrollments equal those of the fall, these institutions would suffer no loss of revenue.

There is no more powerful feedback than revenue. Nothing could facilitate a shift to the Learning Paradigm more swiftly than funding learning and learning-related institutional

outcomes rather than hours of instruction. The initial response to the idea of outcomes-based funding is likely to be "That's not possible." But, of course, it is. As the new paradigm takes hold, forces and possibilities shift and the impossible becomes the rule.

## **NATURE OF ROLES**

With the shift to the Learning Paradigm comes a change in roles for virtually all college employees.

In the Instruction Paradigm, faculty are conceived primarily as disciplinary experts who impart knowledge by lecturing. They are the essential feature of the "instructional delivery system." The Learning Paradigm, on the other hand, conceives of faculty as primarily the designers of learning environments; they study and apply best methods for producing learning and student success.

If the Instruction Paradigm faculty member is an actor-a sage on a stage-then the Learning Paradigm faculty member is an inter-actor-a coach interacting with a team. If the model in the Instruction Paradigm is that of delivering a lecture, then the model in the Learning Paradigm is that of designing and then playing a team game. A coach not only instructs football players, for example, but also designs football practices and the game plan; he participates in the game itself by sending in plays and making other decisions. The new faculty role goes a step further, however, in that faculty not only design game plans but also create new and better "games," ones that generate more and better learning.

Roles under the Learning Paradigm, then, begin to blur. Architects of campus buildings and payroll clerks alike will contribute to and shape the environments that empower student learning. As the role structures of colleges begin to loosen up and as accountability for results (learning) tightens up, organizational control and command structures will change. Teamwork and shared governance over time replace the line governance and independent work of the Instruction Paradigm's hierarchical and competitive organization.

In the Learning Paradigm, as colleges specify learning goals and focus on learning technologies, interdisciplinary (or nondisciplinary) task groups and design teams become a major operating mode. For example, faculty may form a design team to develop a learning experience in which students networked via computers learn to write about selected texts or on a particular theme.

After developing and testing its new learning module, the design team may even be able to let students proceed through it without direct faculty contact except at designated points. Design teams might include a variety of staff: disciplinary experts, information technology experts, a graphic designer, and an assessment professional. Likewise, faculty and staff might form functional teams responsible for a body of learning outcomes for a stated number of students. Such teams could have the freedom that no faculty member has in today's atomized framework, that to organize the learning environment in ways that maximize student learning.

## **MEETING THE CHALLENGE**

Changing paradigms is hard. A paradigm gives a system integrity and allows it to function by identifying what counts as information within the infinite ocean of data in its environment. Data that solve problems that the paradigm identifies as important are information; data that are irrelevant to those problems are simply noise, static. Any system will provide both channels for transmitting information relevant to the system and filters to reduce noise.

Those who want to change the paradigm governing an institution are-from the institution's point of view-people who are listening to the noise and ignoring the information. They appear crazy or out of touch. The quartz watch was invented by the Swiss. But the great Swiss watchmakers responded to the idea of gearless timepieces in essentially the same way that the premiere audience responded to Stravinsky's *The Rite of Spring*. They threw tomatoes. They hooted it off the stage.

The principle also operates in the other direction. From the point of view of those who have adopted a new paradigm, the institution comes to sound like a cacophony-generating machine, a complex and refined device for producing more and louder noise. From the perspective of the governing paradigm, the advocates of the insurgent paradigm seem willing to sacrifice the institution itself for pie-in-the-sky nonsense. But from the perspective of the insurgents, the defenders of the present system are perpetuating a system that no longer works.

But paradigms do change. The Church admits Galileo was right. *The Rite of Spring* has become an old warhorse. Paradigms can even change quickly. Look at your watch.

Paradigms change when the ruling paradigm loses its capacity to solve problems and generate a positive vision of the future. This we very much see today. One early sign of a paradigm shift is an attempt to use the tools and ideas of a new paradigm within the framework provided by the old, or to convey information intelligible in the new paradigm through the channels of the old. This, too, is now happening.

In our experience, people will suffer the turbulence and uncertainty of change if it promises a better way to accomplish work they value. The shift to the Learning Paradigm represents such an opportunity.

The Learning Paradigm doesn't answer all the important questions, of course. What it does do is lead us to a set of new questions and a domain of possible responses. What knowledge, talents, and skills do college graduates need in order to live and work fully? What must they do to master such knowledge, talents, and skills? Are they doing those things? Do students find in our colleges a coherent body of experiences that help them to become competent, capable, and interesting people? Do they understand what they've memorized? Can they act on it? Has the experience of college made our students flexible and adaptable learners, able to thrive in a knowledge society?

How do you begin to move to the new paradigm? Ultimately, changing paradigms means doing everything differently. But we can suggest three areas where changes-even small ones-can create leverage for larger change in the future.

First, you begin by speaking. You begin to speak *within* the new paradigm. As we come to understand the Learning Paradigm, we must make our understanding public. Stop talking about



the "quality of instruction" or the "instructional program." Instead, talk about what it takes to produce "quality learning" and refer to the college's "learning programs." Instead of speaking of "instructional delivery," speak about "learning outcomes."

The primary reason the Instruction Paradigm is so powerful is that it is invisible. Its incoherencies and deficiencies appear as inherent qualities of the world. If we come to see the Instruction Paradigm as a product of our own assumptions and not a force of nature, then we can change it. Only as you begin to experiment with the new language will you realize just how entrenched and invisible the old paradigm is. But as you and your colleagues begin to speak the new language, you will then also begin to think and act out of the new paradigm.

Second, if we begin to talk about the "learning outcomes" of existing programs, we'll experience frustration at our nearly complete ignorance of what those outcomes are-the Learning Paradigm's most important category of information is one about which we know very little now. The place to start the assessment of learning outcomes is in the conventional classroom; from there, let the practice grow to the program and institutional levels. In the Learning Paradigm, the key structure that provides the leverage to change the rest is a system for requiring the specification of learning outcomes and their assessment through processes external to instruction. The more we learn about the outcomes of existing programs, the more rapidly they will change.

Third, we should address the legally entrenched state funding mechanisms that fund institutions on the basis of hours of instruction. This powerful external force severely constrains the kinds of changes that an institution can make. It virtually limits them to changes within classrooms, leaving intact the atomistic one-teacher, one-classroom structure. We need to work to have state legislatures change the funding formulas of public colleges and universities to give institutions the latitude and incentives to develop new structures for learning. Persuading legislators and governors should not be hard; indeed, the idea of funding colleges for results rather than seat time has an inherent political attractiveness. It is hard to see why legislators would resist the concept that taxpayers should pay for what they get out of higher education, and get what they pay for.

Try this thought experiment. Take a team of faculty at any college-at your college-and select a group of students on some coherent principle, any group of students as long as they have something in common. Keep the ratio of faculty to students the same as it already is. Tell the faculty team, "We want you to create a program for these students so that they will improve significantly in the following knowledge and cognitive skills by the end of one year. We will assess them at the beginning and assess them at the end, and we will tell you how we are going to do so. Your task is to produce learning with these students. In doing so, you are not constrained by any of the rules or regulations you have grown accustomed to. You are free to organize the environment in any way you like. The only thing you are required to do is to produce the desired result-student learning."

We have suggested this thought experiment to many college faculty and asked them whether, if given this freedom, they could design a learning environment that would get better results than what they are doing now. So far, no one has answered that question in the negative. Why not do it?

The change that is required to address today's challenges is not vast or difficult or expensive. It is a small thing. But it is a small change that changes everything. Simply ask, how would we do things differently if we put learning first? Then do it.

Those who say it can't be done frequently assert that environments that actually produce learning are too expensive. But this is clearly not true. What we are doing now is too expensive by far. Today, learning is prohibitively expensive in higher education; we simply can't afford it for more and more of our students. This high cost of learning is an artifact of the Instruction Paradigm. It is simply false to say that we cannot afford to give our students the education they deserve. We can, but we will not as long as we allow the Instruction Paradigm to dominate our thinking. The problem is not insoluble. However, to paraphrase Albert Einstein, we cannot solve our problem with the same level of thinking that created it.

Buckminster Fuller used to say that you should never try to change the course of a great ship by applying force to the bow. You shouldn't even try it by applying force to the rudder. Rather you should apply force to the trim-tab. A trim-tab is a little rudder attached to the end of the rudder. A very small force will turn it left, thus moving the big rudder to the right, and the huge ship to the left. The shift to the Learning Paradigm is the trim-tab of the great ship of higher education. It is a shift that changes everything.

---

## CHART I

### COMPARING EDUCATIONAL PARADIGMS

#### **The Instruction Paradigm**

##### **Mission and Purposes**

- Provide/deliver instruction
- Transfer knowledge from faculty to students
- Offer courses and programs
- Improve the quality of instruction
- Achieve access for diverse students

##### **Criteria for Success**

- Learning varies

Inputs, resources

#### **The Learning Paradigm**

##### **Mission and Purposes**

- Produce learning
- Elicit students discovery and construction of knowledge
- Create powerful learning environments
- Improve the quality of learning
- Achieve success for diverse students students

##### **Criteria for Success**

Learning varies

- Quality of entering students
- Curriculum development,expansion
- Quantity and quality of resources
- Enrollment, revenue growth
- Quality of faculty, instruction

### **Teaching/Learning Structures**

- Atomistic; parts prior to whole
- Time held constant,learning varies
- 50-minute lecture,3-unit course
- Classes start/end at same time
- One teacher, one classroom
- Independent disciplines, departments
- Covering material
- End-of-course assessment
- Grading within classes by instructors
- Private assessment
- Degree equals accumulated credit hours

### **Learning Theory**

- Knowledge exists "out there"
- Knowledge comes in chunks and bits; delivered by instructors and gotten by students
- Learning is cumulative and linear

- Learning & student-success outcomes
- Quality of exiting students
- Learning technologies development,
- Quantity and quality of outcomes
- Aggregate learning growth, efficiency

Quality of students, learning

### **Teaching/Learning Structures**

- Holistic; whole prior to parts
- Learning held constant, time varies
- Learning environments
- Environment ready when student is
- Whatever learning experience works
- Cross discipline/department
- Specified learning results
- Pre/during/post assessments
- External evaluations of learning
- Public assessment
- Degree equals demonstrated knowledge and skills

### **Learning Theory**

- Knowledge exists in each person's mind and is shaped by individual experience
- Knowledge is constructed, created,

- Fits the storehouse of knowledge metaphor
- Learning is teacher centered and controlled
- "Live" teacher, "live" students required
- The classroom and learning are competitive and individualistic
- Talent and ability are rare

### **Productivity/Funding**

- Definition of productivity: cost per hour of instruction per student
- Funding for hours of instruction

### **Nature of Roles**

- Faculty are primarily lecturers
- Faculty and students act independently and in isolation
- Teachers classify and sort students
- Staff serve/support faculty and the process of instruction
- Any expert can teach
- Line governance; independent actors

- Learning is a nesting and interacting of frameworks
- Fits learning how to ride a bicycle metaphor
- Learning is student centered & controlled
- "Active" learner required, but not "live" students required
- Learning environments and learning are cooperative, collaborative, & supportive
- Talent and ability are abundant

### **Productivity/Funding**

- Definition of productivity: cost per unit of learning per student
- Funding for learning outcomes

### **Nature of Roles**

- Faculty are primarily designers of learning methods and environments
- Faculty and students work in teams with each other and other staff
- Teachers develop every student's competencies and talents
- All staff are educators who produce student learning and success
- Empowering learning is challenging and complex
- Shared governance; teamwork independent actors

*Change, November/December 1995, pp. 13-25. Reprinted with permission of the Helen Dwight Reid Educational Foundation. Published by Heldref Publications, 1319 Eighteenth St., N.W., Washington, D.C. 20036-1802. Copyright 1995.*

## SUGGESTED READINGS ON ENCOURAGING FACULTY ENGAGEMENT IN ASSESSMENT

- Angelo, T. A. (2000). A vision worth working toward: Assessment in support of learning communities. *Assessment Update*, 12 (2), 3, 5.
- Banta, T. W. (2005, September/October). How much have we learned? *BizEd*. 35-38. Available online at: <http://www.aacsb.edu/publications/Archives/SepOct05/p34-39.pdf>
- Banta, T. W. (2005, November/December). Leaders' views on engaging stakeholders in assessment. *Assessment Update*, 17 (6).
- Banta, T. W., & Lefebvre, L. A. (2001). *Effective practices for academic leaders: Leading change through assessment*. Sterling, VA: Stylus.
- Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. San Francisco, CA: Jossey-Bass.
- Brown, K. (2001). Community college strategies: Why aren't faculty jumping on the assessment bandwagon—and what can be done to encourage their involvement? *Assessment Update*, 13(2), 8-9, 16.
- Brown, F. W., & Moshavi, D. (2002). Herding academic cats: Faculty reactions to transformational and contingent reward leadership by department chairs. *Journal of Leadership Studies*, 8(3), 79-93.
- Burke, J. C. (1999). The assessment anomaly: If everyone's doing it, why isn't more getting done? *Assessment Update*, 11(4), 3, 14-15.
- Cross, K. P., & Steadman, M. H. (1996). *Classroom research: Implementing the scholarship of teaching*. San Francisco, CA: Jossey-Bass.
- Eckel, P., Green, M., & Hill, B. (2001). *On change: Riding the waves of change: Insights from transforming institutions*. Washington: American Council on Education
- Gray, P. J. (1997). Viewing assessment as an innovation: Leadership and the change process. In Gray, P. J., & Banta, T. R. (Eds.), *The campus-level impact of assessment: Progress, problems, and possibilities* (pp. 5-16) (New Directions in Higher Education, Vol. 100).
- Grunwald, H., & Peterson, M. W. (2003). Factors that promote faculty involvement in and satisfaction with institutional and classroom student assessment. *Research in Higher Education*, 44(2), 173-204.
- Huba, M. E., & Freed, J. E. (2000). Shifting the paradigm: Individual and organizational implications. In *Learning-centered assessment on college campuses: Shifting the focus from teaching to learning*. Boston: Allyn & Bacon.
- Hutchings, P., & Shulman, L. S. (1999) The scholarship of teaching: New elaborations, new developments. *Change*, 31(5), 11-15.
- Kotter, J. P. (1995). Leading change: Why transformation efforts fail. *Harvard Business Review*, 73(2), 59-68.
- Leviton, L. C. (2003). Commentary: Engaging the community in evaluation: Bumpy, time consuming, and important. *American Journal of Evaluation*, 24(1), 85-90.
- Loacker, G., & Mentkowski, M. (1993). Creating a culture where assessment improves learning. In T. W. Banta & Associates (Eds.), *Making a difference: Outcomes of a decade of assessment in higher education*. San Francisco, CA: Jossey-Bass.
- Palomba, C. A., & Banta, T. W. (1999). Encouraging involvement in assessment. In *Assessment essentials: Putting principles to work on college campuses* (pp. 53-84). San Francisco, CA: Jossey-Bass.
- Peterson, M. W., & Augustine, C. H. (2000). Organizational practices enhancing the influence of student assessment information in academic decisions. *Research in Higher Education*, 41(1), 21-52.
- Schneider, C. G., & Shoenberg, R. (1999). Habits hard to break: How persistent features of campus life frustrate curricular reform. *Change*, 31(2), 30-35.
- Stufflebeam, D. L. (2002, July 15). Institutionalizing evaluation checklist. Kalamazoo, MI: Western Michigan University, The Evaluation Center. Available online at <http://www.wmich.edu/evalctr/checklists/institutionalizingeval.htm>
- Suskie, L. (2004). Promoting an assessment culture. In *Assessing student learning: A common sense guide* (pp. 34-48). San Francisco: Jossey-Bass Anker Series.
- Suskie, L. (2003). Assessment at Towson University: Lessons learned on keeping assessment thriving. *Assessment Update*, 15(3), 8-9.
- Wergin, J. (2003). *Departments that work: Creating and sustaining cultures of excellence in academic programs*. San Francisco: Jossey-Bass Anker Series.
- Young, C. C., & Knight, M. E. (1993). Providing leadership for organizational change. In T. W. Banta & Associates (Eds.), *Making a difference: Outcomes of a decade of assessment in higher education*. San Francisco, CA: Jossey-Bass.

## SUGGESTED READINGS ON USING ASSESSMENT RESULTS TO IMPROVE TEACHING AND LEARNING

- Angelo, T. A. (1993, April). A "teacher's dozen": Fourteen general, research-based principles for improving higher learning in our classrooms. *AAHE Bulletin*, 45(8), 3-7, 13. Available online: [http://www.csuchico.edu/~lseyder/ceeoc/teachers\\_dozen.pdf](http://www.csuchico.edu/~lseyder/ceeoc/teachers_dozen.pdf)
- Association of American Colleges & Universities. (2002). *Greater expectations: A new vision for learning as a nation goes to college*. Washington, DC: Author. Available online: <http://www.greaterexpectations.org>
- Astin, A. W. (1993). *What matters in college: Four critical years revisited*. San Francisco, CA: Jossey-Bass.
- Bain, K. (2004). *What the best college teachers do*. Cambridge, MA: Harvard University Press.
- Barr, R. B., & Tagg, J. (1995). From teaching to learning: A new paradigm for undergraduate education. *Change*, 27(6), 12-25. Available online: <http://critical.tamucc.edu/~blalock/readings/tch2learn.htm>
- Chickering, A. W., & Gamson, Z. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin*, 39(7), 5-10.
- Chickering, A. W., & Gamson, Z. (1991). Applying the seven principles for good practice in undergraduate education (*New Directions for Teaching and Learning*, No. 47). San Francisco, CA: Jossey-Bass.
- Coalition of Essential Schools. (n.d.). *How to analyze a curriculum unit or project and provide the scaffolding students need to succeed*. Oakland, CA: Author. Available online: [http://www.essentialschools.org/cs/resources/view/ces\\_res/85](http://www.essentialschools.org/cs/resources/view/ces_res/85)
- Gardiner, J. (2002). Research on learning and student development and its implications. In R. Diamond (Ed.), *Field guide to academic leadership*. San Francisco: Jossey-Bass.
- Huba, M. E., & Freed, J. E. (2000). Understanding hallmarks of learner-centered teaching and assessment. In *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning* (pp. 32-64). Needham Heights, MA: Allyn & Bacon.
- Kuh, G. (2001). Assessing what really matters to student learning: Inside the National Survey of Student Engagement. *Change*, 33(3), 10-17, 66.
- Kuh, G. D., Schuh, J. H., Whitt, E. J., & Associates. (1991). *Involving colleges: Successful approaches to fostering student learning and development outside the classroom*. San Francisco, CA: Jossey-Bass.
- Light, R. (2001). *Making the most of college: Students speak their minds*. Cambridge, MA: Harvard University Press.
- McKeachie, W. J., & Others. (2002). *Teaching tips: Strategies, research, and theory for college and university teachers* (11th ed.). Boston, MA: Houghton Mifflin.
- Mentkowski, M., & Associates. (2000). *Learning that lasts: Integrating learning, development, and performance in college and beyond*. San Francisco, CA: Jossey-Bass.
- Miller, M. A. (2007, September/October). Habits die hard. *Change*, vol(no), 7-8.
- Palmer, P. J. (1998). *The courage to teach: Exploring the inner landscape of a teacher's life*. San Francisco, CA: Jossey-Bass.
- Pascarella, E. T. (2001). Identifying excellence in undergraduate education: Are we even close? *Change*, 33(3), 19-23.
- Pascarella, E. T., & Terenzini, P.T. (2005). *How college affects students: A third decade of research*. San Francisco, CA: Jossey-Bass.
- Steen, L. A. (1992, May). 20 questions that deans should ask their mathematics department (or, that a sharp department will ask itself). *AAHE Bulletin*, 44(9), 3-6.
- Strange, C. C., & Banning, J. H. (2000). *Educating by design: Creating campus learning environments that work*. San Francisco: Jossey-Bass.
- Suskie, L. (2004). Using assessment findings effectively and appropriately. In *Assessing student learning: A common sense guide* (pp. 300-317). San Francisco: Jossey-Bass Anker Series.
- Tagg, J. (2003). *The learning paradigm college*. San Francisco: Jossey-Bass Anker Series.

Walvoord, B. E., & Breihan, J. R. (1997). Helping faculty design assignment-centered courses. In D. DeZure (Ed.), *To improve the academy* (Vol. 16) (pp. 349-372). Stillwater, OK: New Forums Press.

Wehlburg, C. M. (2006). *Meaningful course revision: Enhancing academic engagement using student learning data*. San Francisco: Jossey-Bass Anker Series.

What research says about improving undergraduate education. (1996). *AAHE Bulletin*, 48(8), 5-8.

Compiled by Linda Suskie, Middle States Commission on Higher Education  
August 19, 2009

Comparing Educational Paradigms - Barr and Tagg, "From Teaching to Learning", *Change*, Nov/Dec 1995, pp. 13-25

<b>The Instruction Paradigm</b>	<b>The Learning Paradigm</b>
<b>Mission and Purposes</b>	<b>Mission and Purposes</b>
Provide/deliver instruction	Produce learning
Transfer knowledge from faculty to students	Elicit students discovery and construction of knowledge
Offer courses and programs	Create powerful learning environments
Improve the quality of instruction	Improve the quality of learning
Achieve access for diverse students	Achieve success for diverse students
<b>Criteria for Success</b>	<b>Criteria for Success</b>
Learning varies	Learning varies
Inputs, resources	Learning & student-success outcomes
Quality of entering students	Quality of exiting students
Curriculum development, expansion	Learning technologies development
Quantity and quality of resources	Quantity and quality of outcomes
Enrollment, revenue growth	Aggregate learning growth, efficiency
Quality of faculty, instruction	Quality of students, learning
<b>Teaching/Learning Structures</b>	<b>Teaching/Learning Structures</b>
Atomistic; parts prior to whole	Holistic; whole prior to parts
Time held constant, learning varies	Learning held constant, time varies
50-minute lecture, 3-unit course	Learning held constant, time varies
Classes start/end at same time	Environment ready when student is
One teacher, one classroom	Whatever learning experience works
Independent disciplines, departments	Cross discipline/department
Covering material	Specified learning results
End-of-course assessment	Pre/during/post assessments
Grading within classes by instructors	External evaluations of learning
Private assessment	Public assessment
Degree equals accumulated credit hours	Degree equals demonstrated knowledge and skills
<b>Learning Theory</b>	<b>Learning Theory</b>
Knowledge exists "out there"	Knowledge exists in each person's mind and is shaped by individual experience
Knowledge comes in chunks and bits; delivered by instructors and gotten by students	Knowledge is constructed, created,
Learning is cumulative and linear	Learning is a nesting and interacting of frameworks
Fits the storehouse of knowledge metaphor	Fits learning how to ride a bicycle metaphor
Learning is teacher centered and controlled	Learning is student centered & controlled
"Live" teacher, "live" students required	"Active" learner required, but not "live" students required
The classroom and learning are competitive and individualistic	Learning environments and learning are cooperative, collaborative, & supportive
Talent and ability are rare	Talent and ability are abundant



<b>The Instruction Paradigm</b>	<b>The Learning Paradigm</b>
<b>Productivity/Funding</b>	<b>Productivity/Funding</b>
Definition of productivity: cost per hour of instruction per student	Definition of productivity: cost per unit of learning per student
Funding for hours of instruction	Funding for learning outcomes
<b>Nature of Roles</b>	<b>Nature of Roles</b>
Faculty are primarily lecturers	Faculty are primarily designers of learning methods and environments
Faculty and students act independently and in isolation	Faculty and students work in teams with each other and other staff
Teachers classify and sort students	Teachers develop every student's competencies and talents
Staff serve/support faculty and the process of instruction	All staff are educators who produce student learning and success
Any expert can teach	Empowering learning is challenging and complex
Line governance; independent actors	Shared governance; teamwork independent actors

## ***Fair Assessment Practices: Giving Students Equitable Opportunities to Demonstrate Learning***

**Linda Suskie**

**I** am a terrible bowler. On a good night, I break 100. (For those of you who have never bowled, the highest possible score is 300 and a score below 100 is plain awful.) This is a source of great frustration for me. I've taken a bowling class, so I know how I'm supposed to stand and move, hold the ball and release it. Yet despite my best efforts to make my arms and legs move the same way everytime, the ball only rarely rolls where it's supposed to. Why, I wonder, can't my mind make my body perform the way I want it to, every time I roll the ball?

If we can't always control our bodily movements, we certainly can't always control what goes on in our heads. Sometimes we write and speak brilliantly; sometimes we're at a loss for words. Sometimes we have great ideas; sometimes we seem in a mental rut. Is it any wonder, then, that assessment—finding out what our students have learned—is such a challenge? Because of fluctuations in what's going on inside our heads, we inconsistently and imperfectly tell our students what we want them to do. Because of similar fluctuations in what's going on in our students' heads, coupled with cultural differences and the challenges of interpersonal communication, they can't always fully interpret what we've told them as we intended them to, and they can't always accurately communicate to us what they know. We receive their work, but because of the same factors, we can't always interpret accurately what they've given us.

A colleague who's a chemist throws up his hands at all this. Having obtained controlled results in a laboratory, he finds assessment so full of imprecision that, he says, we can never have confidence in our findings. But to me this is what makes assessment so fascinating. The answers aren't there in black and white; we have, instead, a puzzle. We gather clues here and there, and from them try to infer an answer to one of the most important questions that educators face: What have our students truly learned?

### **Seven Steps to Fair Assessment**

If we are to draw reasonably good conclusions about what our students have learned, it is imperative that we make our assessments—and our uses of the results—as fair as possible for as many students as possible. A fair assessment is one in which students are given equitable opportunities to demonstrate what they know (Lam, 1995). Does this mean that all students should be treated exactly the same? No! Equitable assessment means that students are assessed using methods and procedures most appropriate to them. These may vary from one student to the next, depending on the student's prior knowledge, cultural experience, and cognitive style. Creating custom-tailored assessments for each student is, of course, largely impractical, but nevertheless there are steps we can take to make our assessment methods as fair as possible.

**1. *Have clearly stated learning outcomes*** and share them with your students, so they know what you expect from them. Help them understand what your most important goals are. Give them a list of the concepts and skills to be covered on the midterm and the rubric you will use to assess their research project.

**2. *Match your assessment to what you teach*** and vice versa. If you expect your students to demonstrate good writing skills, don't assume that they've entered your course or program with those skills already developed. Explain how you define good writing, and help students develop their skills.

**3. *Use many different measures and many different kinds of measures.*** One of the most troubling trends in education today is the increased use of a high-stakes assessment—often a standardized multiple-choice test—as the sole or primary factor in a significant decision, such as passing a course, graduating, or becoming certified. Given all we know about the inaccuracies of any assessment, how can we say with confidence that someone scoring, say, a 90 is competent and someone scoring an 89 is not? An assessment score should not dictate decisions to us; we should make them, based on our professional judgement as educators, after taking into consideration information from a broad variety of assessments.

Using "many different measures" doesn't mean giving your students eight multiple-choice tests instead of just a midterm and final. We know now that students learn and demonstrate their learning in many different ways. Some learn best by reading and writing, others through collaboration with peers, others through listening, creating a schema or design, or hands-on practice. There is evidence that learning styles may vary by culture (McIntyre, 1996), as different ways of thinking are valued in different cultures (Gonzalez, 1996). Because all assessments favor some learning styles over others, it's important to give students a variety of ways to demonstrate what they've learned.

**4. *Help students learn how to do the assessment task.*** My assignments for student projects can run three single-spaced pages, and I also distribute copies of good projects from past classes. This may seem like overkill, but the quality of my students' work is far higher than when I provided less support.

Students with poor test-taking skills may need your help in preparing for a high-stakes examination; low achievers and those from disadvantaged backgrounds are particularly likely to benefit (Scruggs & Mastropieri, 1995). Performance-based assessments are not necessarily more equitable than tests; disadvantaged students are likely to have been taught through rote memorization, drill, and practice (Badger, 1999). Computer-based assessments, meanwhile, penalize students from schools without an adequate technology infrastructure (Russell & Haney, 2000). The lesson is clear: No matter what kind of assessment you are planning, at least some of your students will need your help in learning the skills needed to succeed.

**5. *Engage and encourage your students.*** The performance of "field-dependent" students, those who tend to think more holistically than analytically, is greatly influenced by

faculty expressions of confidence in their ability (Anderson, 1988). Positive contact with faculty may help students of non-European cultures, in particular, achieve their full potential (Fleming, 1998).

**6. Interpret assessment results appropriately.** There are several approaches to interpreting assessment results; choose those most appropriate for the decision you will be making. One common approach is to compare students against their peers. While this may be an appropriate frame of reference for choosing students for a football team or an honor society, there's often little justification for, say, denying an A to a student solely because 11 percent of the class did better. Often it's more appropriate to base a judgement on a standard: Did the student present compelling evidence? summarize accurately? make justifiable inferences? This standards-based approach is particularly appropriate when the student must meet certain criteria in order to progress to the next course or be certified.

If the course or program is for enrichment and not part of a sequence, it may be appropriate to consider growth as well. Does the student who once hated medieval art now love it, even though she can't always remember names and dates? Does another student, once incapable of writing a coherent argument, now do so passably, even if his performance is not yet up to your usual standards?

**7. Evaluate the outcomes of your assessments.** If your students don't do well on a particular assessment, ask them why. Sometimes your question or prompt isn't clear; sometimes you may find that you simply didn't teach a concept well. Revise your assessment tools, your pedagogy, or both, and your assessments are bound to be fairer the next time that you use them.

## Spreading the Word

Much of this thinking has been with us for decades, yet it is still not being implemented by many faculty and administrators at many institutions. Our challenge, then, is to make the fair and appropriate use of assessments ubiquitous. What can we do to achieve this end?

- *Help other higher education professionals learn about fair assessment practices.* Some doctoral programs offer future faculty studies in pedagogy and assessment; others do not. Encourage your institution to offer professional development opportunities to those faculty and administrators who have not had the opportunity to study teaching, learning, and assessment methods.
- *Encourage disciplinary and other professional organizations to adopt fair assessment practice statements.* A number of organizations have already adopted such statements, which can be used as models. Models include statements adopted by the Center for Academic Integrity (McCabe & Pavela, 1997); the Conference on College Composition and Communication (1995); the Joint Committee on Standards for Educational Evaluation (1994); the Joint Committee on Testing Practices (1988); the National Council on Measurement in Education (1995); and

the first National Symposium on Equity and Educational Testing and Assessment (Linn, 1999); as well as AAHE (1996). (See Assessment Policies, below).

- *Speak out when you see unfair assessment practices.* Call for the validation of assessment tools, particularly those used for high-stakes decisions. Advise sponsors of assessment practices that violate professional standards, and offer to work with them to improve their practices.
- *Help improve our assessment methods.* Sponsor and participate in research that helps create fairer assessment tools and validate existing ones. Collaborate with assessment sponsors to help them improve their assessment tools and practices. Help develop feasible alternatives to high-stakes tests.
- *Help find ways to share what we already know.* Through research, we have already discovered a great deal about how to help students learn and how to assess them optimally. With most of us too busy to read all that's out there, our challenge is finding effective ways to disseminate what has been learned and put research into practice.

As we continue our search for fairness in assessment, we may well be embarking on the most exhilarating stage of our journey. New tools such as rubrics, computer simulations, electronic portfolios, and Richard Haswell's minimal marking system (1983) are giving us exciting, feasible alternatives to traditional paper-and-pencil tests. The individually custom-tailored assessments that seem hopelessly impractical now may soon become a reality. In a generation—maybe less—it's possible that we will see a true revolution in how we assess student learning, with assessments that are fairer for all . . . but only if we all work toward making that possible.

---

*When this article was written, Linda Suskie was director of AAHE's Assessment Forum, and assistant to the president for special projects at Millersville University of Pennsylvania.*

## **Assessment Policies**

Several organizations have developed statements that include references to fair assessment practices. Some are available online:

Code of Fair Testing Practices in Education by the Joint Committee on Testing Practices, National Council on Measurement in Education  
[ericae.net/code.txt](http://ericae.net/code.txt)

Code of Professional Responsibilities in Educational Measurement by the National Council on Measurement in Education  
[www.natd.org/Code\\_of\\_Professional\\_Responsibilities.html](http://www.natd.org/Code_of_Professional_Responsibilities.html)

Leadership Statement of Nine Principles on Equity in Educational Testing and Assessment by the first National Symposium on Equity and Educational Testing, North

Central Regional Educational Laboratory

[www.ncrel.org/sdrs/areas/issues/content/cntareas/math/malnewst.htm](http://www.ncrel.org/sdrs/areas/issues/content/cntareas/math/malnewst.htm)

Nine Principles of Good Practice for Assessing Student Learning by the American Association for Higher Education

[http://www.usc.edu/programs/cet/private/pdfs/9\\_principles\\_assess.pdf](http://www.usc.edu/programs/cet/private/pdfs/9_principles_assess.pdf)

Writing Assessment: A Position Statement by the Conference on College Composition and Communication

<http://www.ncte.org/ccs/12/sub/state6.html>

## References

American Association for Higher Education. (1996, July 25).

Nine principles of good practice for assessing student learning [Online].

Available: <http://www.aahe.org/assessment/principles.htm>

Anderson, J. A. (1988). Cognitive styles and multicultural populations. *Journal Teacher Education*, 24(1), 2-9.

Badger, E. (1999). Finding one's voice: A model for more equitable assessment. In A. L.

Nettles & M. T. Nettles (Eds.), *Measuring up: Challenges minorities face in educational assessment* (pp. 53-69). Boston: Kluwer.

Conference on College Composition and Communication. (1995). Writing assessment: A position statement [Online]. Available: <http://www.ncte.org/ccs/12/sub/state6.html>

Fleming, J. (1998). Correlates of the SAT in minority engineering students: An exploratory study. *Journal of Higher Education*, 69, 89-108.

Gonzalez, V. (1996). Do you believe in intelligence? Sociocultural dimensions of intelligence assessment in majority and minority students. *Educational Horizons*, 75, 45-52.

Haswell, R. H. (1983). Minimal marking. *College English*, 45, 600-604.

Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs* (2nd ed.). Thousand Oaks, CA: Sage.

Joint Committee on Testing Practices. (1988). *Code of fair testing practices in education* [Online]. Available: <http://ericae.net/code.txt>

Lam, T. C. M. (1995). Fairness in performance assessment: *ERIC digest* [Online]. Available: <http://ericae.net/db/edo/ED391982.htm> (ERIC Document Reproduction Service No. ED 391 982)

Linn, R. L. (1999). Validity standards and principles on equity in educational testing and assessment. In A. L. Nettles & M. T. Nettles, (Eds.), *Measuring up: Challenges minorities face in educational assessment* (pp. 13-31). Boston: Kluwer.

McCabe, D. L., & Pavela, G. (1997, December). The principled pursuit of academic integrity. *AAHE Bulletin*, 50(4), 11-12.

McIntyre, T. (1996). Does the way we teach create behavior disorders in culturally different students? *Education and Treatment of Children*, 19, 354-70.

National Council on Measurement in Education. (1995). *Code of professional responsibilities in educational measurement* [Online]. Available: [http://www.natd.org/Code\\_of\\_Professional\\_Responsibilities.html](http://www.natd.org/Code_of_Professional_Responsibilities.html)

Russell, M., & Haney, T. (2000, March 28). Bridging the gap between testing and technology in schools. *Education Policy Analysis Archives* [Online serial], 8(19). Available: <http://epaa.asu.edu/epaa/v8n19.html>

Scruggs, T. E., & Mastropieri, M. A. (1995). *Teaching test-taking skills: Helping students show what they know*. Cambridge MA: Brookline Books.

This article was reprinted from the May AAHE Bulletin (American Association for Higher Education) and published in *Adventures in Assessment*, Volume 14 (Spring 2002), SABES/World Education, Boston, MA, Copyright 2002.

Funding support for the publication of this document on the Web provided in part by the Ohio State Literacy Resource Center as part of the [LINCS Assessment Special Collection](#).

**General Education Assessment Worksheet**

A	B	C	D	E	F	G	H
Middle States competency	Articulate this competency or goal as an assessable learning outcome.	Which Gen Ed requirements (if any) are intended to develop this competency? How?	Are all majors expected to develop this competency? How?	What other requirements (if any) are intended to develop this competency? How?	Do you need to improve student opportunities to develop this competency?	How is this competency now being assessed?	Possible new/improved assessment tools/techniques
Written communication							
Oral communication							
Scientific reasoning							
Quantitative reasoning							
Technological competence							
Critical analysis & reasoning							
Information literacy							
<b>Institutional student learning/development goals:</b>							



## Information on Assessment Models and Best Practices

If you are interesting in finding models and best practices for assessing student learning, here are some resources you can explore:

**Attend a conference.** They are the best places to network with colleagues and learn about the latest assessment practices. Some of the best-known conferences include the following:

- The **Middle States Annual Conference**, held each December, always features presentations and poster sessions on assessment. For more information, visit <http://www.msche.org> and click on Events.
- The **Assessment Institute in Indianapolis**, held in late October or early November, is the country's largest conference devoted solely to assessment in higher education. For more information, visit <http://planning.iupui.edu/conferences/national/nationalconf.html>.
- The **Association for Institutional Research Annual Forum**, held each May, devotes a track to assessment. For more information, visit <http://www.airweb.org>.
- The **National Conference on First Year Assessment**, sponsored by the National Center on the First Year of College and Students in Transition, offers models and best practices that can often be applied beyond the first college year. For more information, visit <http://www.sc.edu/fye/index.html> and click on Events.
- The **Atlantic Assessment Conference** (<http://www.meredith.edu/rpa2009/aac/>), the **Texas A&M Assessment Conference** (<http://assessment.tamu.edu/conference/>), and the **New England Educational Assessment Network Fall Forum** (<http://neean.southernct.edu>) each attracts several hundred people from across the country. Their smaller size makes them ideal for networking.

**Subscribe to *Assessment Update*.** This peer-reviewed quarterly publication is the closest thing the American higher education community has to an assessment journal. It always includes articles on cutting-edge practices by assessment practitioners. To subscribe, visit <http://www.josseybass.com/WileyCDA/WileyTitle/productCd-AU.html>.

**Search the massive *Internet Resources for Higher Education Outcomes Assessment*** sponsored by North Carolina State University (<http://www2.acs.ncsu.edu/UPA/assmt/resource.htm>). Among its hundreds of links are assessment "handbooks," institutional assessment websites, and information on assessment of specific disciplines and skills.

**Join the ASSESS listserv.** This is a fairly low-traffic but useful unmediated listserv in which higher education assessment practitioners share ideas on the nuts and bolts of assessment. To join or search the archives, visit <http://lsv.uky.edu/archives/assess.html>.

**Read key publications.** Among them:

Banta, T. W., Jones, E. A., & Black, K. E. (2009). *Designing effective assessment: Principles and profiles of good practice*. San Francisco: Jossey-Bass.

Bresciani, M. J. (Ed.) (2007). *Assessing student learning in general education: A compilation of good practice case studies*. San Francisco: Jossey-Bass Anker Series.

Lattuca, L. R., Terenzini, P. T., & Volkwein, J. F. (2006). *Engineering change: A study of the impact of ED2000*. Baltimore: ABET, Inc.

## Change



### Counting and Recounting: Assessment and the Quest for Accountability

Lee S. Shulman

From *Change* January/February 2007

*Accounting is essentially a form of narrative.*

Dina Shulman

When my daughter Dina returned from her first class in managerial accounting early in her MBA program, I innocently asked how it had gone. I fully expected her to describe her boredom with the rigors of accounting, since pursuing an MBA was decidedly an afterthought for my iconoclastic daughter, who already held degrees in theatre and social work.

Imagine my surprise when Dina responded that accounting was unexpectedly interesting because, she now realized, it should be understood as a form of narrative, a kind of drama. Within the ethical and technical rules of the field, the task of the accountant is to figure out which of the stories of the company should be told through the medium of its "books." Accounting is basically about creating the plot, characters, and setting of the story. As the instructor explained to the class, "Your task is to render an account: to tell the facts of the case, the story of the condition of a company in an accurate and yet ultimately persuasive way."

I was reminded of this conversation as I read through the successive drafts of the Spellings Commission report, with its persistent refrain that higher education must become more accountable, more transparent, and more open to the scrutiny of its stakeholders. The key word is always "accountability," to which the canonical reaction among educators is a reaffirmation of the remarkable diversity of American colleges and universities and the dangers that accompany the specter of standardized testing and a "one-size-fits-all" approach to assessing the quality of a college education.

In the world of business, an account is a story told in quantitative form. It publicly documents all the income and investments that enter the company and all the products and liabilities that emerge from it, all its assets and debits, all its profits and losses. When the books balance, the account is closed: The story has been told.

Indeed, historian of science Mary Poovey argues in *A History of the Modern Fact* that a significant source for the modern conception of a scientific fact—that which is measurable, replicable, visible, quantitative, and credible—is the invention of double-entry bookkeeping in late-16th century England. Thus accounting was a source for modern scientific conceptions of evidence; then, in full-circle fashion, scientific doctrines became the basis for our contemporary conceptions of account-ability in education.

When I draw our attention, as Dina did mine, to the ghosts of narrative and story-telling that stand behind the counting, measuring, and computations that lie at the heart of modern assessment in the service of accountability, I do not aim to undermine the credibility of assessment. I am not referring to "mere storytelling" as if narrative is a lesser form of discourse. The connections between counting and recounting are built into the etymology of these words in many languages. Thus, in German, to count is *zaehlen* and to tell (a story) is *erzaehlen*. Even in Hebrew, a language with utterly different roots than English or German, the verb for counting is *I'spor*, while the word for telling is *I'saper*.

I believe the lesson is clear. How and what we choose to count and the manner in which we array and display our accounts is a form of narrative—legitimately, necessarily, and inevitably.

#### Tools for Counting and Recounting

When Benjamin Bloom led a group of university examiners in the development of the taxonomy of educational objectives in the late 1940s and early 1950s, their goal was to provide a structure within which assessors could determine which story they wished to tell about the learning of their institution's students. They had determined that most of the instruments then in use to assess students—and thus to render them, their teachers, and their colleges accountable—were exclusively stories of the acquisition and retention of knowledge, of

**Shulman, L., Change, January/February 2007. Vol 39, No. 1**

**<http://www.carnegiefoundation.org/change/sub.asp?key=98&subkey=2169>**

the students' success in recalling facts, events, principles, and concepts they had learned in class or read in their textbooks. Bloom and his colleagues argued that this was an impoverished story, one that missed the most important aspects of the account the examiners needed to give of students' learning.

By elaborating the cognitive outcomes of education into a taxonomy comprised of six categories—ranging from knowledge and comprehension through application, analysis, synthesis, and evaluation—Bloom and his colleagues developed a much richer array of plots and themes for the story of academic performance. A program that appeared to be achieving great success when knowledge alone was measured might look much less impressive if the "higher-order" processes were accounted for. Bloom and his associates also were committed to extending the story from the cognitive to the affective domains in order to include the development of emotions, motivations, passions, and identity.

The power of Bloom's approach to make visible important aspects of learning that would otherwise remain hidden (or to point out their absence) is nicely illustrated by a painful episode in my own history as a learner. When I was an undergraduate at the University of Chicago in the late 1950s, I attempted to cram for the end-of-year comprehensive examination in the history of western civilization—a nine-hour multiple-choice and essay test. I thought I had done quite well on the exam and was thus shocked to receive a "C" for the course. I asked to meet with a member of the Evaluation Office to learn why I had performed so poorly. We sat down and examined my performance, using Bloom's taxonomy as a template. I had "aced" the multiple-choice section, with its emphasis on recall; cramming can be a pretty good strategy for remembering facts and ideas, at least over the short term. But I had simply not studied well enough to integrate the ideas and to be able to synthesize new interpretations and arguments using the knowledge I had crammed into my head.

Had the accounting been limited to a factual knowledge of history, the Shulman narrative would have been one about a highly accomplished student of history. But the richer plot afforded by the design of this assessment told a more complex and less comforting story: Shulman knew the facts of history well but had not yet learned to use them in the service of new ideas or to solve novel problems.

Narratives are enriched not only by changes in plot and theme; introducing new characters as protagonists also has a profound effect. Thus, if the narrative were to examine the learning of discrete sub-groups of students, its complexity and nuance would increase. Is this an institution where students of one particular ethnic background score well across the categories while others do well only in knowledge acquisition but not in the higher-order achievements? Or is this a college where those majoring in the sciences flourish while those studying the humanities flounder? Each of these is a legitimate, "true," and reasonable account—on which the school's accountability will rest. Numbers may offer an illusion of irreducible accuracy and credibility, but they can only be interpreted in the context of the narrative selected and, indeed, the narrative not taken.

The story told by an assessment is thus ultimately a function of the dimensions of measurement that determine the possible directions the narrative might take. So accountability requires that we take responsibility for the story we commit ourselves to telling. We must make public the rationale for choosing that story as opposed to alternative narratives. This requires that we first deliberate with our colleagues and stakeholders about the goals we set, the missions of our schools, and the elaborated conceptions of our purposes.

Only then should we defend the adequacy of the forms of measurement and documentation we employ to warrant the narratives we offer. In the case of educational accountability, we are limited in our recountings by the instruments we use to count. As my colleague Lloyd Bond regularly reminds me, "Since we can't normally measure everything that counts, we can be sure that what will count is what we choose to measure." Taxonomies and indicators are critical aspects of how and with what coherence and credibility these stories can be told.

We can readily see the narrative possibilities for these accounts by examining some of the instruments and indicators that the Spellings Commission singled out. The Collegiate Learning Assessment (CLA) has received a great deal of attention recently and is described in some detail by Richard Shavelson in this issue of *Change*. What story does the CLA tell? The broad domains of its account are critical thinking, analytical reasoning, problem-solving, and writing. The heart of the narrative is the value added by a college education to the educational outcomes of students, rather than the absolute levels they achieve. It chronicles the development of their learning, thinking, judgment, and communication skills and does not aim or claim to assess domain-specific knowledge, skills, values, or appreciations. Thus, students' performance on the CLA does not correlate with their majors. It is currently used to tell a story about institutions, not individual students.

The National Survey of Student Engagement (NSSE) tells a very different kind of story. Although the items are designed to serve as proxies for outcomes, the instrument itself measures the kinds of experiences students have over the course of their academic careers. While the CLA looks for changes in the performance of students, the NSSE is more attuned to the opportunities the institutions offer and the advantage the students take of them. The NSSE describes institutions in terms of their level of academic challenge; the opportunities they provide for active and collaborative learning; the extent and quality of students' interactions with faculty; the availability and access to enriching extra-curricular experiences; and the extent to which the campus offers a supportive environment for learning and student development.

It's no accident that so many institutions (more than 970 for NSSE and 250 for the CLA) have opted to use one or both of these instruments. Each offers a very different narrative of educational opportunities and accomplishments. While they were not designed to fit together elegantly, they do offer different perspectives on this question: What account can be given of this institution's contribution to the education of its students? Notice, however, that neither instrument tells us anything about the discipline-specific aspects of learning. Do students learn to think like historians? Do they learn to reason quantitatively? Do they come to know the fundamental concepts of science and technology that are needed in the 21st century economy?

The Educational Testing Service's Measure of Academic Proficiency and Progress (MAPP), another instrument specifically identified by the Spellings Commission, attempts to tell a story that gets at some of these differences. Its chapter headings are "Reading," "Writing," "Critical Thinking," "Mathematics," "Humanities," "Social Sciences," "Natural Sciences," and—naturally—a total score. But before we leap to the conclusion that in the MAPP we now have a comprehensive, domain-specific map of student learning over time, we must note that the long form of the assessment takes all of two hours and includes 108 items, which is rather sparse for a substantive evaluation across so many areas. And ETS also offers an abbreviated form of the MAPP that contains only 36 items and can be administered in a total of 40 minutes!

My short tour of these tools and instruments (and of course there are many others that could be mentioned) is meant to point up both possibilities and limitations. We are better off with the CLA, the NSSE, and other new tools than without them. **But the bottom line is that the instruments now available for accountability purposes are necessarily short, superficial, and limited.** They are designed to interfere minimally with instruction and to be sufficiently general and unrelated enough to the details of any institution's curriculum that they can be broadly used. **In vivid contrast, the great promise of assessment is its deployment in the service of instruction, its capacity to inform the judgment of faculty and students regarding how they can best advance the quality of learning.** So the challenge before us is to develop systems of assessment and accountability in which the internal uses of assessment for instruction—and the external uses of assessment for accountability and transparency—are carefully weighed. Ultimately, these are the books that need to be balanced—or, when necessary, to be strategically unbalanced.

So what are the lessons to be learned from our sense of accountability as narrative and argument? What tools and approaches can provide the most valid account of the condition of higher education and its constituent institutions? Is the most valid account necessarily the broadest and most comprehensive? Is the best strategy to develop highly specific, narrowly targeted instruments that offer deep insight into particular kinds of learning and development? Should we be looking at institutional performance or at the learning of individual students?

### Seven Pillars of Assessment for Accountability

Most of the principles I want to offer here are familiar, even venerable. The fact that they remain pertinent suggests how persistent many of the challenges of assessment remain.

1. **Become explicit about the story you need to tell and the rationale for choosing it.** An account is one story among the many that could be told about the quality and character of an educational experience. No instrument can claim validity, no account can earn a warrant, without a clear explanation of why this story is being told instead of others. Indeed, it should be clear what the major alternative accounts could be and why they were rejected. Any one form of assessment, however rich, is a compromise, a choice among a set of legitimate possibilities.

2. **Do not think that there is a "bottom line."** An early step in the deployment of any instrument, new or old, should be a process of locating the instrument in a larger conceptual framework that explicitly stipulates what it does measure *and* what it does not. Since there is no real bottom line, the first obligation of the person rendering an account is to take responsibility for locating its unavoidable insufficiencies.

Shavelson does this quite clearly for the CLA in this issue of *Change*, locating its domains of measurement within a figure that sketches out the broader domains that it does not assess. Bloom's classic taxonomies provide tools that can be employed in a similar manner.

Moreover, judgments of validity are never a property of measuring instruments per se. Validity can only be judged when we examine assessment results *in the context of a particular argument or narrative*. The cardinal principle of accountability is that counting is only meaningful and useful in the context of valid recounting. Indeed, we might make a distinction between measurement and assessment in this regard, with assessment referring to the manner in which one arrays, displays, and interprets particular measurements in the service of judgments, decisions, and actions.

3. **Design multiple measures.** As the stakes associated with a measurement rise, the restrictions on its form rise concomitantly—thus the need to move from judgment to measurement and from interpretation to objectivity. But as in any form of social inquiry, the price of precision is narrowness of scope. Therefore, a third principle that follows from the "no-bottom-line" observation is that nearly any use of assessment for serious practical and policy guidance should intentionally employ an array of instruments that will constitute a "union of insufficiencies." It is dangerous to permit highly consequential decisions of policy and practice to rest on the results of a single instrument, however carefully it has been field-tested and ostensibly validated.

In the Texas system of accountability for colleges and universities, for example, more than a dozen instruments are recommended for use, including the National Survey of Student Engagement, the Collegiate Learning Assessment, and multiple indices of access, graduation, and post-graduation success, often broken down by the racial and ethnic backgrounds of the students. Using this array of indicators enables others to render accounts that respond to their questions.

4. **Work on combining multiple measures.** A fourth principle is that a set of instruments, each with its own scores, indices, and observations, will deliver on its promise only if we take on the hard task of developing rules for deciding how to display, organize, and aggregate those indicators for making decisions. Inevitably, those decisions are functions of human judgment—which is, after all, an essential element in any such process, not something to be feared or avoided. On the other hand, there is a good argument to be made for "mechanical combination," in which general policies are debated and determined in such a way that algorithms for systematically

**Shulman, L., Change, January/February 2007. Vol 39, No. 1**

**<http://www.carnegiefoundation.org/change/sub.asp?key=98&subkey=2169>**

combining the available data can be computed objectively. The late Hillel Einhorn of the University of Chicago referred to this process as "expert measurement and mechanical combination."

**5. Remember that high stakes corrupt.** A fifth principle is that high stakes attached to assessments have a tendency to distort the educational and evaluation processes they were intended to support. This is not only because teachers and students are sorely tempted to cheat when the stakes are high. It is also because when test designers know that high stakes are involved, they have a tendency to use items less likely to be uncertain or subject to competing judgments and arguments. As the instruments are weeded of such items or sections, they gain reliability and objectivity but often at the sacrifice of validity and nuance.

The most significant feature of high-stakes assessment is this: The higher the stakes, the greater the likelihood that teachers will teach to the test. These assessments must be designed so that the tests are *worth* teaching to. This is not a trivial challenge. It cries out for a strategy of embeddedness.

**6. Embed assessment into ongoing instruction.** Assess early and assess often. In my early days in Chicago, we used to joke, "Vote early and vote often." High-stakes assessments are likely to be used very late in the course or program where they are employed in the service of accountability. But the later the assessment, the later the knowledge of results, and the less likely it is that the assessments will yield information that can guide instruction and learning. I call these "high-stakes/low-yield" forms of assessment. They may satisfy accountability mavens but have little educative value. Instead, we should develop low-stakes/high-yield forms of assessment, much like the "running records" used by K-12 reading teachers or the routine medical history, physical examinations, or lab tests that physicians and nurses administer.

Assessment should not only serve as an external evaluation and public conscience for higher-education institutions; at the very least, it also should do no harm to instruction, and at best, it should guide, support, and enrich it. When we embed assessment in instruction, it is much more likely that what is assessed will contribute to and be compatible with the core objectives of instruction. If colleges and universities can become active pedagogical laboratories, assessment that is useful for both instruction and accounting will be actively embedded and used continuously.

Embedded measures will necessarily be designed with a different "grain size" from those designed exclusively for external, high-stakes assessments. They will be more particular than general; more dedicated to measuring individual student progress than institutional success; repeatedly administered rather than being single end-of-course events; and highly transparent to students and teachers. They will have quick turn-around times rather than providing the highly secure, secretive, and delayed feedback of current high-stakes environments. This is assessment as a regular physical exam rather than as a public autopsy.

This aspect of assessment emphasizes the need for *bilateral* transparency. That is, the progress students are making needs to be as accessible to them as it is to teachers or policymakers. Such transparency can empower students to take greater control of their own destinies. It is, after all, ultimately the student who must own her or his understanding and progress. Systems of assessment that are opaque, secretive, and slow-responding cripple students' sense of responsibility.

**7. Become an active and collaborative site for research on new forms of assessment, new technologies to support such work, and better strategies for integration of such approaches with instruction.** If the use of single-instrument, high-stakes/low-yield assessment tools will, as many of us have argued over the years, undermine the most important goals and purposes of education, then those of us who design and deploy assessments have a professional and ethical responsibility to design them to contribute more positively to the quality of teaching and learning for all students. The need now is for new assessment research and development, a project that can succeed only if institutions collaborate, experiment, and open their windows so that national work can move our fields ahead.

We need a strategy to combine the local with the national and to meld low-stakes assessment with an accountability approach that will be minimally corrupting. This will require a change in the reward system of higher education to encourage faculty to engage in such experimental approaches to their teaching, rather than worrying that they will be punished if they permit such activity to interfere with more traditional forms of research and scholarship. In the public-policy arena, the culture of competition and ranking, of punitive reactions to honest accounting, of oversimplification via report cards and bottom lines must be resisted.

### **Taking Control of the Narrative**

One of the reasons Dina was so taken with the metaphor of narrative in accounting was that the career she had pursued just before her MBA program was as a psychotherapist. During her graduate study in social work, she had been drawn to "narrative therapy" as an approach to counseling. In narrative therapy, the central idea is that each one of us is living the life of a character in a play or a novel. Some of us feel that we have a great deal of influence over the plot of the play, while others, alas, feel that they are characters in someone else's drama. The goal of the psychotherapy is to support one's clients in seeing the narratives they feel they are living but have no control over, and to develop strategies for becoming the authors of their own stories, able to act responsibly in the situation and exercise real agency over their lives.

I often feel that academics, in the face of the growing volume of calls for accountability, have developed a sense of higher education as victim, swept away by a powerful current over which we can exercise little influence. We think of accountability as a sinister plot invented by others, controlled by the enemy, and designed to take over our professional lives and make us unhappy. We must either paddle upstream, resisting all the way, or just go with the flow, adopting a stance of minimal compliance while hoping to find a little eddy in which we can float about undisturbed. But skilled white-water rafters and canoeists remind us that neither paddling against the

**Shulman, L., Change, January/February 2007. Vol 39, No. 1**

**<http://www.carnegiefoundation.org/change/sub.asp?key=98&subkey=2169>**

current nor going with the flow is a particularly fruitful tactic. The best way to get where *you* want to go when negotiating the rapids in a fast-moving stream is to *paddle faster than the current*.

In this spirit, our responsibility is to take control of the narrative. We educators must take advantage of the deep connections between counting and recounting to define the characters, the plots, the foreground, and the background of the new accountability systems. We must summon the creative energy and ambition to take advantage of the momentum (and resources) unleashed by the new policies and exploit them to initiate the long-overdue progress in assessment needed to improve the quality of learning in higher education.

We are obligated to recount the narratives of most interest to our key stakeholders, but we cannot be limited to those alone. We must display the evidence of teaching and learning (and their embarrassments) through the multiple legitimate narratives we create about our work and our students' fates. We must account for higher-order understanding and critical thinking, in addition to factual knowledge and simple skills. We must tell of the development of civic responsibility and moral courage, even when our stakeholders have not thought to ask for those books.

Moreover, we must make the process through which we render the accounts transparent to our stakeholders. The most important of these stakeholders are our students, who need to feel a sense of agency and responsibility in this relationship as well.

The current quest for accountability creates a precious opportunity for educators to tell the full range of stories about learning and teaching. Counting and recounting can only be pursued together. Counting without narrative is meaningless. Narrative without counting is suspicious. We now have an opportunity to employ the many indicators of learning that we can count for the most important stories we have to tell.

---

*Lee S. Shulman is the eighth president of The Carnegie Foundation for the Advancement of Teaching. He was the first Charles E. Ducommun Professor of Education and professor of psychology (by courtesy) at Stanford University, after being professor of educational psychology and medical education at Michigan State University.*

## Resources

Hillel Einhorn, *Organizational Behavior and Human Performance* 7 (1972): 86-106.

Mary Poovey, *A History of Modern Fact: Problems of Knowledge in the Sciences of Wealth and Society*. Chicago: University of Chicago Press, 1998.



## Middle States Commission on Higher Education

3624 Market Street, Philadelphia, PA 19104-2680

Phone: 267-284-5000 Fax: 215-662-5501 [www.msche.org](http://www.msche.org)

### The Role of Published Tests and Assessments in Higher Education

Published tests and assessments can play an important role in understanding and improving student learning in colleges and universities by adding dimensions and perspectives not available through locally developed tests, rubrics, and surveys. Published tests give colleges a sense of how their students compare against their peers, and some published tests provide detailed feedback that lets colleges easily identify relative strengths and weaknesses in their students' performance. And because published tests and assessments are typically developed by testing professionals, the quality of test questions and problems may be superior to what faculty and staff at individual colleges can develop.

Published tests and assessments are not a panacea, however, for several reasons.

One of the great strengths of American higher education is its diversity. Our country offers community colleges, art schools, engineering schools, theological seminaries, nursing schools, liberal arts colleges, technical institutes, and research universities, to name just a few. Each type of higher education institution appropriately aims to instill a distinct set of knowledge, skills, and competencies in order to prepare its students for successful careers and service to society. While writing and critical thinking skills are important outcomes of any college education, for example, the kinds of writing and critical thinking skills that students need vary according to their aspirations. Students at a culinary school need to learn to write succinctly, while those at a research university need to learn to write extensive, in-depth analyses. Art students need to learn critical thinking skills that emphasize creativity, while business students might find logical reasoning skills more important. Because America's college students have diverse needs and goals, there can be no one test that is appropriate for every college and every program. Published tests and assessments reflect this diversity; the table below gives examples of the variety of writing and critical thinking skills assessed by three higher education instruments.

	<b>Examples of Tested Writing Skills</b>	<b>Examples of Tested Critical Thinking Skills</b>
ETS Measure of Academic Proficiency & Progress (MAPP)	<ul style="list-style-type: none"> <li>• Discriminate between appropriate and inappropriate use of parallelism.</li> <li>• Recognize redundancy.</li> </ul>	<ul style="list-style-type: none"> <li>• Evaluate competing causal explanations.</li> <li>• Determine the relevance of information for evaluating an argument or conclusion.</li> </ul>
ACT Collegiate Assessment of Academic Proficiency (CAAP)	<ul style="list-style-type: none"> <li>• Formulate an assertion about a given issue.</li> <li>• Organize and connect major ideas.</li> </ul>	<ul style="list-style-type: none"> <li>• Generalize and apply information beyond the immediate context.</li> <li>• Make appropriate comparisons.</li> </ul>
Council for Aid to Education Collegiate Learning Assessment (CLA)	<ul style="list-style-type: none"> <li>• Support ideas with relevant reasons and examples.</li> <li>• Sustain a coherent discussion.</li> </ul>	<ul style="list-style-type: none"> <li>• Deal with inadequate, ambiguous, and/or conflicting information.</li> <li>• Spot deceptions and holes in the arguments made by others.</li> </ul>

Some other tests and assessments, meanwhile, are not intended to measure college-level learning and are thus inappropriate for use at this level. The National Assessment of Adult Literacy, for example, tests the abilities to understand materials such as job applications, transportation schedules, and food labels and to perform arithmetic computations such as balancing a checkbook or figuring a tip—all important skills that are typically (and appropriately) taught at the basic (through grade 12) rather than higher education level.

Concern with published tests and assessments is that, unless students have compelling incentives to give them their best effort, the results will not accurately reflect what students have truly learned. While students have a clear incentive to do their best on certification and licensure examinations, it can be difficult to motivate them to do their best on other published tests. Developing compelling yet ethical incentives is a challenge; because all tests and assessments have inherent imperfections, it is inappropriate to make any single test a “gatekeeper” on which a certain score must be earned in order to, say, pass a course or earn a degree.

Yet another concern with published tests and assessments available for higher education is that they often have more limited evidence of their quality than published tests used in basic education. While validation studies at the K-12 level can involve tens of thousands of students, studies of higher education tests often involve far smaller numbers of students from institutions that may not be a representative sample of all colleges and universities. While test publishers continue to work diligently to research and document the validity and reliability of their tests, at this time we cannot have the same level of confidence in higher education test results that we have at the K-12 level.

A final concern with some published tests is that they do not yield enough useful feedback to help colleges identify specific shortcomings and make necessary improvements. The Collegiate Learning Assessment, for example, yields only one global score reflecting a plethora of skills such as understanding data in tables and figures, marshaling evidence from different sources, and distinguishing rational from emotional arguments. Without specific feedback on student performance on each of these skills, colleges whose students perform poorly on the CLA have no idea which skills their students lack and cannot address deficiencies without further research. Indeed, the brochure “CLA In Context 2004-2005” notes, “To use the CLA in a diagnostic manner, you will need to combine the CLA results with other data you collect.” Many colleges may not wish to risk diminishing the quality of their students’ education by diverting scarce resources from the essential business of teaching and learning to costly research studies.

Published tests and assessments *can* yield valuable insight into student learning at the higher education level, but only if (1) they correspond to the college’s goals for student learning, (2) they yield useful feedback that will help the college identify areas that need improvement, (3) they have convincing evidence of their quality (validity and reliability), and (4) students have compelling incentives to give the tests their best effort. Because there is no one perfect instrument, published tests and assessments should only be used in combination with other evidence of student learning, including locally-developed measures, job placement rates, and the like, in order to draw a more accurate overall picture of student learning.

All who are concerned with the future of American higher education can take steps to ensure that students graduate with appropriate knowledge, skills, and competencies. First, we can continue to support the American system of accreditation, which requires all accredited colleges to provide clear, compelling, and appropriate evidence of rigorous student achievement. Second, we can continue to value the rich diversity of American higher education and acknowledge that no one test can adequately evaluate the knowledge, skills, and competencies expected of all of America’s college students. Finally, we can encourage the development and use of assessment tools appropriate to each field of study and each sector of American higher education, so that all students graduate fully prepared for successful careers and productive service to society.

*Linda Suskie, Middle States Commission on Higher Education  
March 25, 2006*





# STUDENT LEARNING ASSESSMENT

*Options and Resources*



---

**Middle States Commission on Higher Education**

---

STUDENT  
LEARNING  
ASSESSMENT

Options and Resources

SECOND EDITION

Published by the  
Middle States Commission on Higher Education  
3624 Market Street  
Philadelphia, PA 19104  
Telephone: 267-284-5000  
Fax: 215-662-5501  
[www.msche.org](http://www.msche.org)

© 2007, Copyright by the Middle States Commission on Higher Education  
Second Edition  
All rights reserved.

First Edition, 2003: This publication replaces *Framework for Outcomes Assessment* (1996) and *Outcomes Assessment Plans: Guidelines for Developing Assessment Plans at Colleges and Universities* (1998).

Second Edition, 2007: The publication *Assessing Student Learning and Institutional Effectiveness: Understanding Middle States Expectations* (2005) replaces Chapter 4 in the First Edition.

Reprinted August 2008 with typographical error corrections.

Permission is granted to colleges and universities within the jurisdiction of the Middle States Commission on Higher Education to photocopy these guidelines for the purpose of institutional self-study and continuing institutional improvement. Bound copies may be purchased through the publications order form, which is available on the Commission's website at [www.msche.org](http://www.msche.org).

*Printed in the United States of America*

# Contents

<b>Acknowledgements</b> . . . . .	v
<b>Introduction</b> . . . . .	1
How Accreditation Helps, Accreditation Standards; Purpose and Scope of this Handbook; Guiding Principles; The Organization of this Handbook; and MSCHE Website Resources	
<b>1. Motivating and Involving Campus Communities</b> . . . . .	5
Partners in Teaching and Learning; and Leading Assessment Initiatives	
<b>2. Learning Goals</b> . . . . .	10
Benefits of Clearly Articulated Learning Goals; Relationship among Learning Goals at All Levels; First Steps Towards Developing Learning Goals; Defining Learning Goals before Selecting Assessment Methods; Ensuring the Quality and Relevance of Learning Goal Statements; and Resources for Creating Student Learning Goal Statements	
<b>3. Evaluating Student Learning</b> . . . . .	27
Formative and Summative Assessment; Benchmarking; Direct and Indirect Methods for Assessing Student Learning; Quantitative vs. Qualitative Evidence; Other Methodological Considerations; Key Questions When Choosing Assessment Instruments; Easy-to-Implement Tools and Techniques; and Other Assessment Tools	
<b>4. Understanding Middle States Expectations for Assessing Student Learning and Institutional Effectiveness</b> . . . . .	54
What is the Assessment of Institutional Effectiveness (Standard 7)? What is the Assessment of Student Learning (Standard 14)? Why does the Commission Expect Student Learning and Institutional Effectiveness to be Assessed? What are the Characteristics of Assessment Processes that Meet Middle States Expectations? What Should Institutions Document Regarding Assessment? How Should This Information be Organized and Formatted for Review by the Commission and Its Representatives? How are the Documentation of Institutional Assessment and Student Learning Assessment Related? What Might the Commission and Its Representatives Look For in Assessment Documentation?	
<b>5. Using Results to Improve Teaching and Learning.</b> . . . . .	59
Institutional Support Strategies Designed to Encourage the Use of Assessment Results; Translating Assessment Results into Better Learning; and A Message to Faculty: The Interconnectedness of Assessment, Teaching, and Learning.	

**Appendices:**

1. Assessment Standards in <i>Characteristics of Excellence in Higher Education</i> . . . . .	75
2. Enhancing the Campus Climate for Assessment: Questions for Academic Leaders . . . . .	84
3. Assessment Practices Quiz . . . . .	86
4. Key to “Assessment Practices Quiz” . . . . .	87
5. Department/Program Student Outcomes Survey . . . . .	88
6. Learning Goals and Assessment Techniques . . . . .	91
7. From Effect to Cause: A Brainstorming Exercise. . . . .	92
8. Student Learning Styles: Frequently Asked Questions . . . . .	93
<b>References</b> . . . . .	97

## List of Figures

1. Relationship between Institutional and Program Goals . . . . .	13
2. Relationship between Program and Course Goals. . . . .	14
3. Leading Questions for Developing Learning Goals . . . . .	22
4. Teaching Goals Inventory . . . . .	23
5. Examples of Direct and Indirect Measures of Student Learning (Course, Program, and Institutional Levels). . . . .	29
6. Criteria-based Rating Scales (Rubrics) . . . . .	43
7. Excerpt from a Simple Rating Scale. . . . .	44
8. Example of a Detailed Rating Scale . . . . .	46
9. Example of a Holistic Scoring Guide (For Critical Thinking). . . . .	48
10. Student Self-reflection Questions for a Course or Program . . . . .	49
11. Example of a Test Blueprint . . . . .	50
12. Commonly-administered Measures of Critical Thinking . . . . .	52
13. Considerations when Deciding to Use Portfolios . . . . .	53
14. Using Assessment Results at the Institutional Level: Maintaining Mission and Achieving Vision. . . . .	64
15. Using Assessment Results at the Program Level: Preparing Students for Future Success . . . . .	66
16. Using Assessment Results at the Course Level: Ensuring Learning . . . . .	68
17. Strategies to Improve Student Learning . . . . .	71

# Acknowledgements

The Middle States Commission on Higher Education is grateful for the assistance of its 2001 Advisory Panel on Student Learning and Assessment for developing the basic text for this publication, with additional contributions by the Commission, its staff, and some colleagues in higher education.

## The Advisory Panel on Student Learning and Assessment

[Positions as of Date Appointed: June 2001]

**Dr. Peter J. Gray (Chair)**, Associate Director, Center for Support of Teaching and Learning, Syracuse University

**Dr. Michael J. Dooris**, Director, Planning Research & Assessment, The Pennsylvania State University

**Dr. Thomas V. Fernandez**, Professor of Physics, Nassau Community College

**Dr. Bruce Keith**, Assistant Dean for Academic Assessment, United States Military Academy

**Dr. Armand S. LaPotin**, Professor of History and Academic Program Coordinator, SUNY College at Oneonta

**Dr. Elizabeth Larsen**, Professor, Coordinator of General Education, West Chester University of Pennsylvania

**Dr. Jodi H. Levine**, Assistant Vice Provost for University Studies, Temple University

**Dr. Rosalyn Lindner**, Senior Advisor to Provost for Assessment, SUNY College at Buffalo

**Dr. Peter J. Miller**, Assistant Professor, Department of Physical Therapy, University of the Sciences in Philadelphia

**Dr. Paula E. Peinovich**, Vice President for Academic Affairs, Excelsior College

**Ms. Linda A. Suskie**, Director of Assessment, Towson University; Former MSCHE Fellow

**Dr. Barbara E. Walvoord**, Director, Kaneb Center for Teaching and Learning; Concurrent Professor of English; and Fellow, Institute for Educational Initiatives, University of Notre Dame

**Prof. Katrina A. Zalatan**, Assistant Professor of Management, Hartwick College

## Additional Support Provided by:

**Dr. MaryAnn Baenninger**, *First Edition Editor-in-Chief*; MSCHE Executive Associate Director

**Ms. Siobhan Underwood**, MSCHE Graduate Fellow and 2001 Assistant to the Advisory Panel

**Ms. Linda Suskie**, *Second Edition Editor-in-Chief*; *MSCHE Vice President*

**Ms. Jean Avnet Morse**, *Editor*; MSCHE President

**Mr. Oswald M. T. Ratteray**, *Editor*; MSCHE Associate Director for Communication



# Introduction

Colleges and universities have long defined and assessed student learning using course-embedded assessments of student learning, such as tests, papers, projects, as well as standardized or “custom” qualitative and quantitative measures. All of these are valid and valuable assessment tools if used properly.

In order to reach out more effectively to students and to the public, the Middle States Commission on Higher Education revised its accreditation standards, *Characteristics of Excellence in Higher Education: Eligibility Requirements and Standards for Accreditation*, to refine the requirements and recommendations for establishing learning goals and assessing student achievement.

The members also concluded that the process of defining and assessing student learning would assist faculty in their teaching, students in selecting institutions and in managing their own learning, and institutions in planning and supporting students.

This handbook serves as a resource for institutions seeking a bridge between the Commission’s standards for accreditation and the practical daily challenges of assessment and continuous improvement.

## *How Accreditation Helps*

It is in the interests of currently enrolled and prospective college students, faculty members, parents, high school teachers and guidance counselors, legislators, employers, and the general public to be informed consumers of higher education.

One function of accreditation is to provide the public with an explanation of the broad scope of higher education and to assure the public that the goals of higher education have been achieved by evaluating each institution within the context of its mission.

One of the means by which the public can understand higher education is through information about the assessment of student learning. As an institutional accreditor, the Middle States Commission on Higher Education, with the support of its institutional members, acts on the judgments of volunteer peer reviewers who certify that institutions assess themselves in all areas, including student learning.

## *Accreditation Standards*

Among the principles that guided the revision of the Commission’s standards is greater emphasis on institutional assessment and the assessment of student learning. By complying with the standards, accredited institutions assure the public that they provide quality higher education. **Specifically, the Commission’s process demonstrates that institutions identify student learning goals for educational offerings that are appropriate to its higher education mission; that its offerings display appropriate academic content, rigor, and coherence; that its curricula are designed so that students demonstrate college-level proficiency in general education and essential skills, including oral and written communication, scientific and quantitative reasoning, critical analysis and reasoning, technological competence, and information literacy; and that assessment demonstrates that students at graduation have achieved appropriate higher education goals.**

The accreditation standards relating to assessment are intended to foster and cultivate the progress of member institutions. They are not intended to be prescriptive. Each standard stresses the significance of self-study and peer review as a developmental activity. The ways in which individual institutions carry out assessment activities and determine the extent to which their goals for student learning have been met is an institutional prerogative. Because of the diversity of institutional types, missions, and educational practices that characterize the members in the Middle States region, *Characteristics*



provides institutions with guidance on how different types of institutions might fulfill each standard.

*Characteristics of Excellence in Higher Education* reflects this renewed and increased emphasis on institutional assessment and the assessment of student learning in several ways. Each of the 14 standards is accompanied by “fundamental elements” that guide the institution in assessing itself on the standard.

With the increasing use of alternative forms of delivery, including distance learning and asynchronous delivery, a focus on outcomes becomes even more essential. Student learning outcomes form a “common currency” with which one can judge the equivalence and value of various learning experiences.

The standards are organized into two subsections: Institutional Context and Educational Effectiveness. The concluding standards of each of these two subsections require that an institution define, evaluate, and continually refine its overall goals (Standard 7), with special emphasis on goals for student learning (Standard 14):

**Standard 7: The institution has developed and implemented an assessment plan and process that evaluates its overall effectiveness in achieving its mission and goals, its efficiency in the use of its resources, and its effectiveness in assuring that its students and graduates achieve the appropriate learning and other outcomes.**

**Standard 14: Assessment of student learning demonstrates that the institution’s students have knowledge, skills, and competencies consistent with institutional goals and that students at graduation have achieved appropriate higher education goals.**

These standards are mutually supportive, because they recognize the centrality of student learning to institutional effectiveness and stress that the assessment of outcomes should be integrated into the institutional planning process. See Appendix 1 for an expanded description of these standards.

## ***Purpose and Scope of this Handbook***

This handbook is intended to clarify principles and methods for setting goals for student learning within the context of institutional mission, for using methods chosen by the institution for evaluating the achievement of these goals, and for using the information gathered to continue to improve student learning. **It is not an expansion of the Standards for Accreditation described in *Characteristics*; it is meant only as a resource.**

Teams that evaluate institutions at the time of their decennial self-studies and evaluators who review institutions’ Periodic Review Reports, Follow-up Reports, and Substantive Change proposals will use the standards themselves, rather than this handbook, to assess the institution.

The audience for this handbook includes all stakeholders of a college or university, including faculty, students, staff, administrators, and the general public. It is intended to assist directly those responsible for setting goals for student learning and for evaluating the achievement of those goals.

This handbook describes:

- How faculty and staff members can define clearly student learning and affective goals appropriate for an institution’s mission;
- Various direct and indirect methods of evaluating student learning and the value and appropriate use of each approach;
- How student learning can be improved by relating outcomes to the institution’s operations and resources; and
- How traditional methods of teaching and learning can be enhanced to produce clear and useful information about how and what students are learning, both inside and outside the classroom, so that faculty, students, the institution, and the general public can benefit from improvements.

The handbook presents various possible means of meeting the Commission’s standards. It describes various contexts and options for assessing student learning, and it provides resources as examples of how institutions *might* approach the assessment of student learning on their campuses. It also discusses

some of the many considerations that should be explored before intensive institutional effort is directed at articulating learning goals, choosing means of evaluating the accomplishment of those goals, and crafting an institutional plan for assessment.

## ***Guiding Principles***

This handbook serves as a starting point for institutions beginning or enhancing their self-assessment activities, particularly those activities related to student learning.

It is written for faculty, staff, and administrators—those who will actually be leading and conducting assessment efforts on their campuses. Its purpose is not limited to providing a route to achieving accreditation or reaffirmation of accreditation, but rather it is intended as a resource and guidebook for institutional self-reflection, improvement, and achievement of the best possible outcomes for students.

It also can be a resource for those who wish to learn about assessment practice in general: what it is, why it is important, who benefits from it, how it is accomplished, and how accreditation supports assessment.

Six guiding principles serve as the framework for this handbook, and they are relied on throughout the handbook. These principles are adapted from AAHE's principles for good practice in assessing student learning (Astin, 1991) and *Assessment in practice: Putting principles to work on college campuses* (Banta, Lund, Black, and Oblander, 1996).

While the recommendations, ideas, resources, and perspectives in this handbook are offered as examples and as flexible models and blueprints, there is no “one size fits all” type of institutional assessment or student outcomes assessment. Thus, the principles presented here should serve as the basis and guiding structure for assessment activities and the resources described in the handbook should serve as possible tools—among many—for achieving institutional goals.

### **Guiding Principle 1: Existing Culture**

Begin by acknowledging the existence of assessment throughout the institution in order to ensure that the assessment plan is grounded in the institutional culture.

### **Guiding Principle 2: Realistic Plan with Appropriate Investment of Resources**

Plans for assessment at the program, school, and institutional levels should be realistic and supported by the appropriate investment of institutional resources.

### **Guiding Principle 3: Involvement of Faculty and Students**

Academic leadership is necessary in order to gain the support and involvement of faculty members, staff, administrators, and students across the institution.

### **Guiding Principle 4: Clear Goals**

Assessment activities should be focused by clear statements of expected student learning (knowledge, skills, and competencies).

### **Guiding Principle 5: Appropriate Methods**

Assessment should involve the systematic and thorough collection of direct and indirect evidence of student learning, at multiple points in time and in various situations, using a variety of qualitative and quantitative evaluation methods that are embedded in courses, programs, and overall institutional processes.

### **Guiding Principle 6: Useful Data**

Data gained through assessment activities should be meaningful. They should be used, first, to enhance student learning at the institutional, program, and course levels; second, in institutional planning and resource allocation; and third, to evaluate periodically the assessment process itself for its comprehensiveness and efficacy.

The guiding principles are intended to help institutions answer the following general questions:

- ❑ What should our students learn? (Chapter 2)
- ❑ What are our institutional strengths and challenges for improvement? (Chapters 1 and 2)
- ❑ How are we currently organized for evaluating learning? (Chapters 3, 4, and 5)
- ❑ What activities have we conducted to define and evaluate all of our institutional goals, with special emphasis on goals for student learning? (Chapters 3 and 4)
- ❑ What existing evidence do we have regarding student learning and achievement, and what have we learned from that evidence? (Chapters 3 and 5)
- ❑ What actions will we take to build on our strengths and to address our weaknesses regarding student learning? (Chapter 5)

Each of the chapters in this handbook focuses on a different component of the assessment process and describes considerations, options, and resources related to that component. The chapters are meant to stand alone or in combination with each other, and the handbook can be valuable to support different approaches.

### ***The Organization of This Handbook***

Readers of this handbook will be approaching the task of evaluating student learning from many vantage points. Some institutions will have existing institution-wide assessment plans that need refinement or adaptation. Other readers will be embarking on cyclic evaluation of their assessment plans at the course, program, or institutional level. Still other readers will be just starting to help their institutions set goals for student learning at the course, program, and institutional level.

In order to address the needs of every institution and every reader, this handbook starts with the development of learning goals and individual assessment strategies, builds to the creation of a written assessment plan, and ends with a chapter on using assessment results.

Even readers from institutions that have an existing plan can benefit from the focus in this handbook on setting learning goals and assessing the related outcomes.

# 1

## Motivating and Involving The Campus Community

**T**he purpose of assessment is to engage a campus community collectively in a systematic and continuing process to create shared learning goals and to enhance learning. Those who have direct instructional and supportive contact with the students and those who lead assessment initiatives are responsible for motivating and involving the rest of the campus community.

The best way to motivate the community is to promote an understanding of the benefits that assessment brings to students, faculty, the institution, and the public. The extent to which learning goals and assessment processes that already exist will form the core of a more clear and integrated assessment process.

Students, of course, want to attend the institution that suits them best. Parents want the best value, or perhaps the “cultural capital,” that an institution affords. Parents and students are interested in which institution will provide them with the education that will result in a job, or acceptance to graduate or professional school. Employers are interested in the “product” that a college or university produces.

Faculty members, too, have a vested interest in students being informed about their choice of a college or university to attend. It is much easier, and more enjoyable, for faculty members to teach students who are appropriately prepared for their courses, either through earlier preparation or through foundation courses at the institution.

### *Partners in Teaching and Learning*

All campus members are partners in teaching and learning and have a role in evaluating and enhancing student learning. Those who have direct instructional and supportive contact with students include faculty, library and information literacy professionals, and student support service professionals.

### **Faculty and Students**

Faculty members traditionally have had the primary responsibility for facilitating student learning. They determine what students should learn, both across the curriculum and within individual courses or programs, and how students should demonstrate their learning. Faculty members devise methods of gathering evidence of student learning and collaborate with other faculty members in evaluating student learning in their majors and academic programs. They use this information to create a true partnership of learners with their students and to improve student learning.

Huba and Freed (2000) provide examples of how to develop this partnership and the benefits it offers.

Faculty members, who are trained as disciplinary experts, as scholars, and as researchers, can amplify their skills by exploring further how students learn best.

First-year students arrive at their institutions eager to embark on collecting the credits toward the credential they believe they need to ensure their long-term economic futures. If they are traditional residential students, they also seek to experience college life in its totality. Returning students or continuing education students may have other, even more pragmatic, reasons for attending college. Often, however, neither group has had sufficient prior experience in reflecting on how to learn, evaluating the extent of what they have learned and what they still need to discover, or using their new knowledge and skills.

These faculty and students are partners who find themselves engaged together in the pursuit of knowledge, skills, and affective development. They must cooperate to produce the best learning possible.

Sometimes, however, faculty members engage in this process automatically, without questioning the tacit assumptions underlying their concept of teaching and learning.

For instance, end of semester term papers are regularly assigned as a means of evaluating student learning. Yet the typical term paper assignment is a good illustration of how traditional approaches may not necessarily effectively foster learning.

The paper may be assigned early in the semester, without the requirement of an outline or draft of the paper during the course of the semester. The professor's concept of what an excellent paper should be, including its form, its content, and the process for completing it, may not have been communicated effectively to the student when the paper was assigned. Therefore, the student may not have engaged in a process designed to meet those expectations. Furthermore, the paper may be graded after the semester is officially finished, may contain no comments, and may or may not be returned to the student.

Assessment of student learning is not a means of decreasing the autonomy of faculty members. It is a means of increasing the mutual engagement of faculty members, staff, and students in providing an optimal learning experience.

It is important for committed faculty members and other institutional leaders to focus on the faculty, staff, and student partnership, and to avoid top-down or prescriptive rules for accomplishing assessment. For example, it may be tempting for institutions to pressure faculty members to orient professional programs too heavily towards goals that promote only practical skills. Conversely, the need to define student learning goals might be perceived as a potential threat to academic freedom. Close partnership of faculty, librarians, student affairs professionals, and students in defining learning goals consistent with institutional mission should avoid such extremes.

### **Library and Information Literacy Professionals**

Not all learning occurs in the classroom. Therefore, library and information literacy professionals also have a critical role in the process of enhancing student learning. Together with faculty, students, and other staff members, they can address the full range of learning in a student's college career. Ideally, methods of facilitating student learning should exist in other divisions of an institution or should be integrated into coursework. A focus on information literacy is an important component in achieving this objective, especially if it is integrated into curricular and co-curricular facets of the institution.

The information literacy paradigm consists of five skills for learners (Association of College and Research Libraries, 2000):

- Determining the nature and extent of needed information;
- Accessing information effectively and efficiently;
- Evaluating critically the sources and content of the information being sought, and incorporating selected information in the learner's knowledge base and value system;
- Using information effectively to accomplish a specific purpose; and
- Understanding the economic, legal, and social issues surrounding the use of information and information technology, as well as observing laws, regulations, and

institutional policies related to the access and use of information.

The principles of information literacy are invoked any time a student attempts to learn anything in any discipline. To the extent that the expected learning involves the use of resources available in or through the library, librarians and faculty share responsibility for various aspects of the process for teaching information literacy.

The subject of information literacy is explained in detail in a separate Commission publication of guidelines on information literacy (Middle States Commission on Higher Education, 2003) and clarified even further in Standards 11 and 12 of *Characteristics of Excellence*, (2006).

### **Student Support Service Professionals**

Those who administer services such as residential life, advising, career development, learning support, service learning, and financial aid are partners with faculty members and students in developing outcomes and assessing student learning. For example, student development personnel help students to develop their own ethical values and to achieve the institution's goal of graduating students who are responsible citizens in a multicultural society.

### ***Leading Assessment Initiatives***

Effective leadership is necessary to create a culture that values student learning assessment within the institutional context.

### **Campus Leaders**

Leaders should involve all constituencies in understanding how assessment can be helpful, in identifying the learning goals that are most important to the community, in assessing outcomes, and in using the results.

Campus leaders might sponsor faculty-led discussions of the issues and concerns related to assessment, present workshops led by internal or external experts, organize faculty and student forums that provide an overview of assessment on campus, address larger assessment issues, and answer assessment questions.

Faculty and staff members who work with students should have ownership of the assessment process as well as full and continuing administrative support for its implementation. Such support is best gained through the public recognition of faculty and staff members' ongoing efforts and accomplishments in assessing student learning. Clear recognition demonstrates an institution's commitment to a culture that values the enhancement of student learning.

Palomba and Banta (1999) note that one or more leaders should take responsibility for leading the campus-wide assessment process, that resources must be committed, and that the institutional priority of assessment should be explicit. Activities to involve faculty, staff, and students should be sponsored by the academic leadership and, in particular, supported by the chief academic officer.

Appendix 2 lists some self-reflection questions for chief academic officers to help them gauge their own level of commitment to assessment activities and to reflect on how they might enhance the campus climate for assessment. Appendix 3 is a short "quiz" that can be used on campuses to stimulate discussion about assessment. Although the quiz is designed for true or false responses, many of the questions are more ambiguous than they appear in order to generate a dialogue about assessment; it is not a test of assessment knowledge. Explanations of the "correct" answers may be found in Appendix 4.

In addition to campus conversations, the institution's leaders, and particularly the president and the provost, can be partners with faculty to introduce and to establish assessment by adopting the following approaches:

- ❑ Increase the awareness of assessment on campus, articulate and define assessment issues and priorities, and identify the institution as an environment that supports assessment practices.
- ❑ Acknowledge assessment activities that already exist and promote fuller participation in assessment activities by facilitating communication and discussion among the institution's members, with the goal of achieving shared responsibility for assessment.

- ❑ Be a sponsor of assessment who shares leadership in bringing about this change in the campus culture.
- ❑ Bring participating members into the assessment process by identifying existing coalitions and developing new coalitions, with the goal of opening the process to as many as possible.
- ❑ Provide funding and other incentives for participation in the assessment process, institutionalizing assessment, and integrating assessment into the faculty and staff roles and rewards process.
- ❑ Provide a clear charge to an appropriate campus assessment committee responsible for communicating expectations for assessment.

### **Institutional Context**

The institutional context is grounded in the institution’s mission, and it is shaped by the institutional culture.

**Mission.** An institution’s mission, at both broad and specific levels, serves as the context within which to assess student learning, and it is important that mission serves as the backdrop for assessment efforts at the institutional, program, and course levels. An institution’s broad contexts will shape overall goals for student learning and how that learning is demonstrated.

For instance, a comprehensive university in the public sector will have a mission driven in large part by the needs and interests of the state, while a private comprehensive university’s mission may focus on the interests of its founders or trustees. In the case of the public university, accountability to the state may include the demonstration of service to the community, success in workforce development, and the ability to keep intellectual resources within the state, in addition to the demonstration that students are liberally educated. Private colleges may have missions that focus solely on the liberal arts.

A comprehensive public sector institution may articulate assessment goals through an emphasis on natural resource management, mining or agronomy programs, characteristic of that state’s economy.

A private comprehensive university, such as a faith-based institution, may develop assessment goals related to an ecclesiastical mission.

Within the context of its mission and broad context, each college or university will have subsidiary and more specific purposes. Thus, one university might stress the development of civic leadership or technical expertise, and another institution might stress pre-professional development and global citizenry.

The link between mission and the development of goals—in this case, goals for student learning—is clearly expressed in Standard 1 of *Characteristics*, which requires that an “institution’s stated goals and objectives, consistent with the aspirations and expectations of higher education, clearly specify how the institution will fulfill its mission.” For example, training excellent teachers or insightful managers may express one institution’s mission, while training academic scholars and researchers are desired goals for another institution.

Failure to clarify the institution’s goals and strengths may result in misallocated resources and confused students and applicants. Students may come to an institution, for instance, with the goal of becoming nurses or biomedical technicians, only to find that the health sciences do not fit well within the institution’s mission. The result is that they may find a “disconnect” between the mission and goals of a college or university and the learning outcomes that its students hope to achieve for themselves.

Leaders of institutional assessment initiatives, then, should refer back constantly to the institutional mission and should articulate to faculty, administrators, board members, and donors the fundamental importance of designing learning goals that are consistent with the institutional mission in order to serve both their students and faculty.

**Institutional Culture.** As an institution begins to plan for assessment, it is important that it consider the particular aspects of institutional culture that might affect the form and process of assessment practice on its campus. Respect for how consensus is achieved and recognition of existing structures, both official and unofficial, will pave the way for reaching the ultimate goal of improving student learning. Following are some questions that leaders

or potential leaders of campus assessment initiatives can ask themselves as they embark on new or changed assessment activities:

- What is the quality of communication on campus? Can it be improved before implementing an assessment plan?
- How is decision-making handled on campus, both formally, and informally?
- What is the level of trust on campus? If trust is a problem, how can it be earned?
- What person or persons on campus are perceived to hold unofficial power? How can those persons be convinced of the benefits of assessment? How can they serve as sources of support for assessment initiatives?
- What is the system of apportioning resources on campus? If there are concerns about equity of resource distribution (or perceived concerns), can they be addressed before implementing an assessment plan?
- What is the process by which curricula and programs are approved and revised? Does the process present any impediments to assessment, and can it be improved or streamlined?
- Are there collective bargaining agreements, union-wide or local, which could either support or impede assessment practices? For example, are union funds available to support faculty development that could be used for assessment? Do union rules restrict “official” work during the summer?

Leaders of assessment on campus should consider these and other factors that could influence the institution’s ability to do assessment well. If the campus culture is not functioning in a manner that is likely to be supportive of assessment, it would be useful to take steps to “heal” the culture before instituting large-scale assessment initiatives. Schein (1996) offers some excellent advice on assessing campus culture and suggests that it may be beneficial to an institution to enlist the advice of an external consultant. A consultant can review and

describe the campus culture, bring objectivity to the review, and offer a fresh perspective. This consultant could work with the chief academic officer or other campus leader to articulate the underlying assumptions of the culture, reveal any “surprises” lurking beneath the surface, and devise strategies for change.



# 2

## Learning Goals

**G**oals for student learning are the foundation of meaningful assessment. Statements of desired student outcomes can be derived through a variety of effective methods at the institutional, program, and course levels. This chapter describes the benefits of having clearly articulated learning goals, explores the characteristics of learning goal statements, and provides resources for implementing participatory processes for developing goals.

### *Benefits of Clearly Articulated Learning Goals*

Clearly articulated statements of what each institution expects its students to learn at the course, program, and institutional levels are important to students, faculty, staff, the institution, and the public for many reasons.

#### **Benefits for Students**

Statements of student learning goals benefit students because they:

- Explain the sometimes “hidden agenda” (e.g., the expectation that students analyze relationships between causes and effects, rather than simply learn substantive material)
- Prioritize which goals are most important
- Provide assurance that a student has not “missed” an important goal

- Help students to understand the nature of skills acquired for use in other contexts—during and after college

#### **Benefits for Faculty and Staff**

Statements of student learning goals benefit faculty and staff because they:

- Identify what to teach, including discipline-specific knowledge and skills, as well as the discipline’s perspective and values
- Provide structure for co-curricular programs
- Determine what will be evaluated at the conclusion of the course or program
- Ensure that skills that should be taught throughout the curriculum actually are included in instruction and evaluation of specific courses

#### **Benefits to the Institution**

Statements of student learning goals benefit the institution because they:

- Publicize to the institution’s constituents evidence that it can demonstrate the accomplishment of clearly-defined student learning goals
- Ensure that goals the institution itself values are assessed, rather than those used by external assessors (e.g., sophisticated analytical math skills versus minimal national competency levels)

- ❑ Ensure that student learning outcomes are suited to the mission of the institution
- ❑ Ensure that core institutional values (e.g., professional career development and approaches of different cultural perspectives) are sufficiently incorporated
- ❑ Ensure that general education skills, such as proficiency in oral and written communication, the ability to think critically and analytically, and the ability to be effective decision-makers and problem-solvers are included in programmatic plans
- ❑ Ensure that the personal growth and affective development of students are addressed
- ❑ Focus attention on the use of direct methods of assessing student learning, supported by meaningful indirect methods, instead of potentially less meaningful indirect measures often used by:
  - external assessors (e.g., graduation rates, cost efficiency, etc.)
  - internal assessors (e.g., student evaluations of faculty)

### **Benefits for the Public**

Statements of student learning goals benefit the public because they:

- ❑ Enable students to choose an institution based on a particular mission
- ❑ Satisfy accountability needs of legislators, funding agencies, and others
- ❑ Help the public to understand more clearly what an institution seeks to accomplish

### ***Relationship among Learning Goals At All Levels***

Before developing or revising learning goals institution-wide, it is important to consider the relationship among learning goals at the institutional, program, and course levels. In addition, different institutions might develop

goals and assess them at varying levels, depending on the needs of the institution.

### **Learning goals at the institutional, program, and course levels**

Students learn specific content and skills in each course. In aggregate, those courses, together with other program experiences such as academic advising, internships, and faculty-directed research by students, should result in the desired student outcomes at the program level. Similarly, goals at the program level combine with general education goals, extra- and co-curricular goals, information literacy goals, and other goals (for example, ethical and civil leadership goals) to create institutional goals. In other words, goals at the institution, program, and course (or activity) levels are interconnected, complimentary, and reciprocal.

Institutions differ in the way that they characterize the relationship between general education goals and institutional goals. In one model, the institution develops a set of overall institutional learning goals stemming from its mission; these goals serve as the super-ordinate (highest level) goals from which program and course level goals flow. In this format, general education goals are essentially programmatic goals; that is, the general education program is one of the programs whose goals contribute to the achievement of overall institutional goals.

In another model, the institution adopts general education goals as overall institutional goals. In this approach, academic and co-curricular program goals would contribute to the achievement of the umbrella-like general education goals, which are essentially institutional goals.

Standard 14 of *Characteristics*, the Assessment of Student Learning, includes language that is most similar to the first model presented above—that is, it describes the assessment of student learning at the institutional, program, and course levels. Standard 12, General Education, also includes its own “fundamental element” related to the assessment of general education.

The Commission, however, is not concerned with the language that an institution uses to describe various levels of learning goals, nor is it concerned with the specific type of hierarchical structure an

institution adopts when defining its goals. It is concerned that the institution develops a coherent set of goals, that those goals stem from the institutional mission, and that goals at the subordinate levels contribute to the attainment of goals at the higher levels. The way in which a particular institution defines general education goals relative to institutional goals depends on the institution's mission (e.g., a specialized institution is unlikely to adopt general education goals as institutional goals). It also depends on how the general education program is structured (e.g., Is it "modular" or are its goals met in part through the major? Are some of its goals met through student affairs programs?) Finally, developing general education goals depends on the institution's history and culture.

Figure 1 and Figure 2 describe levels of learning goals at the institutional, program, and course levels. Although some institutions actually create matrices like these to aid them in formulating goals, this matrix is not presented as a model of how goals should be illustrated but, rather, as an abstraction to help the reader understand the relationships between levels of learning goals.

**Institutional and Program Goals.** Figure 1 illustrates hypothetical relationships among learning goals or statements of student outcomes at the institutional and program levels. The desired outcomes at the institutional level provide the outline or framework for connecting goals at the program level into a coherent whole.

These illustrations include goals for disciplinary and major programs, general education, and a student-affairs oriented program, such as residence life. Goals from various activities and initiatives contribute to overall student affairs goals. Because student affairs and academic programs both contribute to the overall education of the student, goals from each of these programs work together to fulfill institutional goals.

**Program and Course Goals.** Figure 2 illustrates how program goals provide a framework for course goals. It presents general education as a program, but one could well imagine the institutional goals cited here as general education goals instead. Notice also that some of the goals for programs

overlap and that not all programs work toward meeting all institutional goals.

Figure 2 depicts the hypothetical relationship between a subset of the program goals presented in Figure 1 and sample goals from courses in each of those programs. Notice, for instance, that one of the goals in the course World Art History, to "identify and analyze major works representing several different cultures," contributes to the general education program goal to "recognize and appreciate artistic and literary contributions of diverse cultures," which in turn contributes to the institutional goal to prepare "global citizens."

At the course level, the faculty member teaching World Art History will have many additional goals. Some will contribute further to the achievement of general education goals, but others may contribute to the achievement of goals for the major in Art History. Because of the interdependence among goals and the course and program levels, it could be impractical, or perhaps impossible, to specify all of the links between goals for each course and for each program in matrices. It is more important to strive for conceptual coherence, ensuring that learning goals at the various levels are understandable, meaningful, and accessible to faculty and students alike.

## Flexibility

In Standard 14, the Commission recognizes that institutions will be "flexible in their approach to defining student learning goals at these different levels, such as repeating goals (some general education goals, for example) across programs or defining the goals at the institutional or program level as synthesis of the goals set at the program and course levels."

For instance, an institutional goal for undergraduate students to become proficient in information literacy may be achieved through a combination of the active involvement of faculty in information literacy instruction (see Chapter 1), a first-year introduction to library and learning resources presented by a librarian, required assignments in a general education course, and/or a series of substantial research papers required in the major. The goals for student learning in each of these three situations, when combined, may fulfill the

Figure 1

## Relationship between Institutional and Program Goals

*Note:* Not all programs are designed meet all institutional goals. Some cells are left blank for illustrative purposes only, not to imply that these goals cannot be met by the programs used in this example.

Institutional-level Goals	Programs and Program-level Goals				
	General Education	Residence Life	Business	History	Chemistry
<b>Leadership</b>	Function effectively as a team member to produce a scholarly product	Develop leadership skills  Apply conflict resolution skills in a living-learning environment	Develop leadership potential in self and others	Analyze historical perspectives on individual and political leadership	
<b>Global Citizens</b>	Recognize and appreciate artistic and literary contributions of diverse cultures  Exhibit engaged citizenry and value community service	Develop an appreciation for cultural and ethnic diversity	Value and exhibit comfort with cultural differences in business practices  Function effectively as a team member to run a small business	Recognize and value culturally-diverse historical perspectives	Demonstrate an ability to work as a team with a diverse group of students
<b>Technologically Sophisticated Individuals</b>	Use technology effectively to communicate and analyze information		Use technology effectively to communicate and analyze information related to business	Use technology effectively to communicate and analyze information related to history	Use technology effectively to collect, analyze, and display data
<b>Effective Communicators</b>	Write and speak proficiently	Communicate effectively in social situations		Communicate effectively, orally, and in writing about historical topics	Cogently present research data and analyses in written, visual, and oral formats
<b>Critical Thinkers</b>	Distinguish critical from non-critical information		Use critical thinking to analyze business case studies	Critically analyze historical events and trends using scholarly techniques	Apply critical thinking skills to design an experiment that tests an hypothesis  Collect, analyze, and interpret data relevant to testing an hypothesis

Figure 2

## Relationship between Program and Course Goals

Program	Program Goal	Course or Activity	Course Goal
<b>General Education</b>	Recognize and appreciate artistic and literary contributions of diverse cultures	World Art History	Identify and analyze major works representing several different cultures.
		Caribbean Literature	Demonstrate familiarity with themes and genres of classic and contemporary Caribbean literature.
<b>Residence Life</b>	Apply conflict resolution skills in a living-learning environment	First Year Student Orientation Program	Work effectively as part of a group to analyze and resolve a hypothetical interpersonal conflict.
		Seminar led by resident assistants	Develop a plan, in cooperation with floor-mates, for handling conflicts as they arise.
<b>Business Administration</b>	Function effectively as a team member to run a small business	Introduction to Marketing	Develop a feasible marketing plan for a small business.
		Capstone in Business Administration	Work with a team of students to develop, plan, manage, and market a small business.
<b>History</b>	Communicate orally and in writing about historical topics	Modern American History	Present a cogent oral analysis of one long-term effect of the Cold War.
		Medieval History	Present a cogently-written, critical analysis of gender and class roles in Medieval England.
<b>Chemistry</b>	Collect, analyze, and interpret data relevant to testing an hypothesis	Introductory Chemistry Laboratory	Replicate chemical reactions using appropriate laboratory techniques.
		Introductory Biochemistry	Collect, analyze, and interpret data relevant to an hypothesis supplied by the instructor.

institutional goal. Thus, an institution need not articulate specific means of achieving a particular goal at the institutional level, or assess it at that level, if it has chosen to assess it at the course and program levels.

Learning goals are discussed throughout the remaining three chapters of this handbook. Chapter 3 discusses the means by which the attainment of goals is assessed. Chapter 4 describes Middle States expectations for assessing student learning and institutional effectiveness. Chapter 5 focuses on using assessment results and how information about whether or not goals have been attained can be used to adapt curricula and programs.

### ***First Steps Towards Developing Learning Goals***

The process of developing learning goals should begin with a “situation audit” or inventory of what exists and which practices have been successful. Practices that are identified will provide information for developing a plan for the assessment of student learning, establishing goals, and identifying assessment measures.

#### **The Situation Audit: Taking an Inventory and Starting with What Already Exists<sup>1</sup>**

A basic tenet for evaluating student learning is to begin with successful assessment activities already in place. Whether the objective is to develop learning goals and assessment techniques for an individual course, an entire program, or the institution as a whole, an inventory of what exists provides a strong foundation for later success.

An excellent method of gauging the level of an institution’s existing evaluation of student learning is to survey the assessment practices embedded at the course, program, and institutional levels. Peter Ewell has referred to this as a “situation audit”—i.e., an inventory of information already on hand that may provide evidence of student learning.

<sup>1</sup> This topic is placed here in the section on goals so that it appears early in the book, and it is referenced again in later chapters. Its placement also emphasizes the point that institutions should examine existing learning goals and develop new ones before making decisions about adopting previously-used measures.

Angelo, Ewell, and Lopez (2001) recommend that institutions begin assessment by “rounding up information you already have.”

**Institutional Level.** At the institutional level, an audit may be accomplished easily by cataloging the means used to assess the entire student body through the activities of offices of institutional research, student affairs, career services, the library, and information management. Most institutions have existing information from some or all of the following:

- Surveys of student satisfaction and engagement that are designed and administered nationally and locally
- Alumni career and satisfaction surveys
- Tests: standardized and/or locally-created
- Statistics, such as placement and retention rates
- Program reviews of both academic and support programs
- Reports by instructional librarians on information literacy and collaboration with faculty members

**Program Level.** At many institutions, each department and program institutes evaluations of its students that are independent from those of other departments and programs. The choice of instruments and assessment activities is often idiosyncratic, grounded in the approach that is typical of each discipline. A comprehensive and well-designed institution-wide checklist of possible types of assessment activities may help each department to create an accurate description of its assessment activities.

At the program level, the checklist for assessment activities might include:

- Senior capstone theses, papers, individual or group projects, and performances or other presentations
- Student portfolios

- Student research participation
- Departmental student and alumni surveys
- Standardized tests of subject area or broad skills
- Reports from student internship supervisors

Additional assessment activities may be suggested by disciplinary accreditors who issue guidelines and standards for intended student learning, required or suggested educational experiences, recommended evaluation methods, and expectations for the use of results.

A survey designed to document assessment practices at the department or program level can assist departments in identifying where there are gaps in the learning goals they are assessing, duplicative teaching efforts, and the usefulness of existing assessment results.

Such an inventory also can offer departments a basis for comparing themselves with other departments, as well as creating an institution-wide resource of where to find advice about instituting assessment on their own campuses. Appendix 5 is an example of such a survey.

**Course Level.** The commitment of individual faculty and teams of faculty is essential. Reviewing existing course-based assessment practices can help faculty members to reflect on assessment practices that have become routine. A review of course materials can provide useful insights into what students may or may not be learning.

A well-constructed course-level checklist might include:

- Embedded assessment elements faculty prepare, such as syllabi, curricula, instructional materials and methods, assignments, exams, and quizzes
- Direct evidence of student learning and development, such as student products and performances resulting from embedded assignments, tests, and other educational experiences
- Indirect indicators such as surveys, placement, and other institutional research data. These indicators can provide both

qualitative and quantitative information over time and across situations.

A more thorough discussion of course-embedded assessment techniques is presented in Chapter 3 of this handbook. It describes the relative uses of quantitative and qualitative information, as well as direct and indirect methods of evaluating student learning.

### **Examining Existing Practices for Success**

Angelo, Ewell, and Lopez (2001) advocate building assessment plans and practices from those activities on campus that are already successful. When a “situation audit” of course-based, programmatic and/or institution-wide assessment is complete, local best practices will surface as models for additional assessment initiatives.

Faculty members and students probably already have a good sense of what is working best on a campus. For example, there may be anecdotal evidence that graduates of one program have particularly strong research skills, while students in another program may be especially adept at using and adapting what they have learned to solve unforeseen problems while working as interns. An audit of teaching and assessment practices used by successful programs will produce models for other departments.

Ideally, some of the faculty members or departments that have been evaluating student learning will have used the results of the evaluation to change practices and to enhance student learning. These efforts also can motivate and guide others in the institution.

Data collected from a comprehensive audit can be used to answer critical questions about which existing assessment practices on campus can form the core of the institution’s assessment program and to identify the most critical gaps for which new assessment techniques are needed. Perhaps the most important benefit of conducting a situation audit is that the data gathered become a foundation for developing learning goals.

---

---

## Starting with Successful Programs

**Identifying successful programs and courses early can help later when organized assessment is started. Starting with the assessment of successful programs offers several benefits:**

- **Effective teaching/learning efforts of faculty members and students are validated.**
- **The admissions, public relations, and development offices have substantive research information to use when publicizing the institution and its programs.**
- **External stakeholders have concrete, rather than anecdotal, evidence of the quality of the institution and its programs.**
- **Faculty members and administrators in other programs can learn from the successes of their colleagues.**

---

---

## *Defining Learning Goals before Selecting Assessment Methods*

The most important step in developing successful methods for evaluating student learning is to develop meaningful, clear, and realistic goals for student learning at the course, program, or institutional level. These goals or statements of expected student learning are different from the actual evidence or the data gleaned from evaluations of student learning. Goals are the basis for determining how best to collect, assess, and interpret the data in order to improve. Data collection not tailored to goals will not provide information about the achievement of desired student learning, nor will it lead to new approaches to teaching and learning.

The goals or statements of student learning are hypotheses for what qualities or attributes will

characterize students after they have completed a course or program, or after they have graduated from the college or the university.

The data generated from actual tests, surveys, or instruments used to gauge the outcome of the educational experience are the actual assessments<sup>2</sup>.

For example, the learning goal might be to develop analytical skill. After a student has taken a course intended to promote this skill, he or she should have better analytical skill. This abstraction, analytical skill—the quality or attribute that students should possess after taking the course—is a generalized notion of what should be achieved. To evaluate the achievement of analytical skill, a test of learning might include problems that can be solved with syllogisms. The syllogistic problems “operationally define”—i.e., make concrete the abstraction “analytical skill.” Thus, success in solving the problems (as indicated by scores above the norm, scores that surpass pretest scores, or scores that differ from those of similar students who did not take the same course) would indicate success in acquiring analytical skill.

Another way to move from the abstract to the specific when articulating student learning goals is to state those goals in terms of what, specifically, a student should be able to do in order to demonstrate that desired learning has occurred. In other words, what observable student behaviors should result from a learning experience? For example, astronomy faculty members may agree that their students will understand basic concepts about the solar system, but they may have differing opinions about what constitutes “basic concepts” and what it means to “understand” them. Do “basic concepts” refer to basic facts about each planet or also to theories about how the solar system was created? Should students memorize those basic facts, or should they be able to use information about our solar system to speculate about the characteristics of other solar systems?

It is important, therefore, to understand that the qualities or attributes that students should exhibit after a learning experience should be operationally

---

<sup>2</sup> In research terms, educational experiences are the “independent variable or treatment,” the assessments are the methods, and their results would be called the “dependent variable.” The student learning outcomes, then, *depend* upon the educational experiences.



defined in order to be assessed meaningfully. For example, a statement of student learning (a learning goal) might be that a student will think critically after completing an introductory course in philosophy. Another learning goal might be that, after completing a service-learning course, a student have greater appreciation for others who are different. Each of these goals can be operationally defined, and then learning can be documented by a test or other instrument created to assess the specific goal. The results of the assessment demonstrate (or do not) the outcome one would expect to see—i.e., What would a student’s performance on this particular assessment look like if he or she is a critical thinker? What would the student’s performance look like if he or she is a person with appreciation for differences between people?

Appendix 6 is a worksheet for an exercise that faculty and staff members can use to begin to develop learning goals for courses and programs and to begin to think about how those goals might be achieved, how they might be assessed, and how a course or program might be altered to ensure greater student learning. The worksheet contains space for only three goals, in order to emphasize that the focus should be on important goals.

The remainder of this chapter 2 is devoted to developing learning goals; Chapter 3 is devoted to evaluating those goals.

## ***Ensuring the Quality and Relevance of Learning Goal Statements***

The institution can ensure the quality and relevance of learning goal statements by focusing on those that are most important, widely accepted by the various stakeholders, meaningful, sufficiently explicit, and interconnected among the various academic levels and curricula within the institution.

### **Key Learning Outcomes**

Effective statements of expected student learning are focused on the most important goals of a course, program, or institution. They are not a collective list of goals that are idiosyncratic to a few faculty or staff members. Attempts to evaluate every possible goal can overwhelm an institution

with tasks, provide too much information, and dilute the focus on areas that need the most attention.

The departmental, school, or institutional mission statement, as described in *Characteristics* (Standard 1), should provide the basis for determining the most important goals at each level. It is useful to concentrate statements of expected learning outcomes by asking, “What are the *most* important learning outcomes we seek for our students in the context of the goals of our institution/program?” For example, the programs and learning outcomes of an institution whose mission includes giving each student a strong spiritual grounding may emphasize different learning outcomes from those of an institution whose mission includes teaching its students technical skills.

### **Widely Agreed-upon Concepts**

Statements of expected learning outcomes will not be effective unless they are developed collaboratively and widely accepted by stakeholders: faculty members, students, employers, alumni, and others affected by or concerned with the program or institution. While it is unlikely that there will be unanimous agreement on expected learning outcomes, there should be a shared sense among most members regarding which learning is most important. The mission of the institution and the subsidiary missions of departments and programs serve as the natural sources for shared expectations.

### **Communication of Learning Goals**

If the institutional community shares learning goals and if they are expressed clearly, then the resulting statements of expected learning outcomes can be used by the entire campus community. Clearly-expressed expectations for the learning outcomes of courses and programs can help students to focus their studies and, as a result, to learn more effectively. Prospective students who are aware of the specific types of expected learning outcomes of a program to which they are considering applying can make a better-informed decision about whether the program meets their

needs, especially when evidence is available that those goals actually are achieved.

Faculty members who teach prerequisite courses or “service” courses can prepare students better for later courses and programs if they are familiar with the expected learning outcomes of subsequent courses or courses in the target program.

For example, faculty members in the English department who are familiar with the expected learning outcomes of the theater department’s programs and courses can better meet the needs of theater students taking literature courses, and physics faculty members can meet the needs of engineering students.

### **Meaningful Learning Goal Statements That Lead to Improvement**

Meaningful statements of student learning goals address learning as a multidimensional and integrated process, occurring over time. They do not focus on trivial learning outcomes. Stated cogently and clearly, meaningful learning goals will lead to the improvement of teaching and learning at the course, program, and institutional levels. The importance of each learning goal should be obvious to students, faculty, and prospective employers.

Meaningful learning goals stress generalizable and higher-order thinking skills rather than memorization of facts or very simple conceptual understanding. For example, a goal to identify grammatical forms (past participles, etc.) is, in most cases, not as meaningful as a goal of being able to write and speak grammatically. Similarly, the successful memorization of important historical dates is not as meaningful as a goal for students to be able to place historical events within a social and political context, to draw meaningful comparisons between events, and to analyze current events within an historical framework. For both of these examples of more meaningful or higher-order goals, the more trivial goals of memorizing dates and acquiring the names of parts of speech probably will be achieved naturally in the course of achieving the larger goal.

### **Sufficiently Explicit Learning Goals**

Although it is not helpful for statements of student learning goals to be so specific that they focus on unimportant or trivial outcomes, it is important for statements to be sufficiently *explicit* for all stakeholders to have a common understanding of their meaning.

For instance, one goal for an undergraduate psychology program might be for students to exhibit proficiency in conducting research. While faculty members may implicitly understand what this goal might mean, increasing the specificity of the goal would enhance its clarity and allow for more direct assessment of the attainment of the goal. For example, a statement of the goal might read: “Students will learn the statistical, organizational, writing, and analytical skills necessary to conduct meaningful and valid scientific research.” Statements then could describe the evidence needed to demonstrate that students have achieved the knowledge and abilities related to each of these components.

### **Interconnectedness of Student Learning Goals**

Student learning can occur at many levels and in many venues:

- Course, program, and institutional levels (Standard 14);
- Undergraduate, graduate, and professional program levels (Standard 11);
- General education curricula (Standard 12);
- Related educational activities, such as basic skills, certificate programs, experiential learning, non-credit offerings, and distance or distributed learning (Standard 13); and
- In co-curricular and extracurricular activities (Standard 9).

An institution’s curriculum may address particular learning outcomes in different complementary or overlapping courses and programs. Statements of learning outcomes for courses or programs should recognize and clarify these relationships, and student learning outcomes assessment plans should be structured to avoid duplication.

---

---

## Choosing Learning Goals

### Start with success

**Determine which learning goals are already being assessed and what data may be available to assess other goals**

- Institutional Level
- Program Level
- Course Level

### Ensure relevance of goals

- Identify key learning outcomes
- Use widely agreed-upon concepts
- Communicate goals
- Select important and meaningful goals
- Be explicit
- Integrate goals in different areas and levels

### Choose goals that can lead to improvement

### Emphasize higher-order thinking skills

**Define learning goals before choosing assessment methods**

- Operationally define each goal
  - Tailor data collection to defined goals
- 
- 

## *Resources for Creating Student Learning Goal Statements*

This section includes a discussion of several specific resources for crafting actual statements of student learning goals. Each of these resources presents opportunities both for “brainstorming” and for comprehensively and systematically reviewing sets of possible goals.

### Existing Learning Goals

An institution already may have explicit and clear statements of student learning goals in place, and it is important to resist the urge to redesign entire sets of course, program, or institutional goals if they already exist. The focus should be on developing those that are missing, those that are unclear, those that have changed, or those that are complementary. For instance, many institutions developed learning goals for general education programs when those programs were initially created, and some may have been substantially revised during subsequent program review. However, faculty members may determine that no measures are being used to assess whether the goals have been achieved. The goals, then, may not need to be re-written; instead, they should be evaluated for their importance and relevance, and they should be supplemented with additional goals where appropriate. Of course, it is still important to measure student outcomes in these areas.

### Existing Syllabi and Course Descriptions

Existing syllabi and catalogue descriptions provide a logical starting point for developing learning goals at the course level, because many faculty members already include learning goals in their syllabi, and many course descriptions include statements about course content, goals, and/or what the student should be able to do once the course is completed. Existing goals such as these can be reviewed for their relevance to programmatic mission, and faculty members should be encouraged to think about whether these goals reflect the current or changing focus of their disciplines. Examples of syllabi that already address learning goals can serve as resources for faculty members who have not previously

developed goals, and faculty members familiar with the process can serve as mentors.

### **Leading Questions**

Leading questions also can serve as the basis for a brainstorming exercise in developing learning goals.

The leading questions listed in Figure 3 can be tailored to apply to any discipline and can be refined to address more specific outcomes. In general, they focus attention on the most important learning goals for individual programs.

The following additional questions related to basic improvement are adapted from the work of Stufflebeam (2001):

- What alternative educational practices and experiences are available, and what are their comparative advantages over current practices at our institution?
- What are the characteristics of a good syllabus that can serve as a guide to teaching and learning?
- What facilities, materials, and equipment are needed to ensure success in reaching our educational objectives?
- What are the roles of faculty members, students, and others in the pursuit of learning?
- Is the course providing learning experiences for all of the students who are enrolled?
- Is a particular course worth the required institutional investment?
- Is the course meeting the minimum accreditation requirements for the discipline?
- Is the course equal to or better than analogous courses at comparable institutions?

### **Analysis of Student Work**

The work products and performances of students that are the result of existing assignments and tests may already embody the important learning outcomes, even though the faculty member may not have explicitly conceived or stated those goals. A retrospective review can suggest how the original statements of goals might be revised, especially if the unstated learning goals are important and if the evaluation methods are valid.

For instance, a political history course may not have the explicit goal of increasing awareness of and participation in current political events. Nevertheless, after taking the course, students may report increased civic awareness, an increased tendency to vote, and increased participation in local political activity. Therefore, it may make sense to make the development of political awareness an explicit course goal and to revise the course accordingly.

### **Inventory of Teaching Goals**

Faculty members and students can use the Teaching Goals Inventory (Angelo and Cross, 1993), shown in Figure 4, to identify the priority of various learning goals in courses and programs. For example, individuals or groups of faculty members could use the inventory to understand better their expectations for a single course, a course with many sections, or a series of courses. They can use the inventory to review faculty documents, to examine existing disciplinary accreditation guidelines and standards, and to analyze direct evidence of student learning and development. Students also could complete the inventory so that the institution can determine whether students and faculty share the same perceptions about the relative priority of different types of learning.

Figure 3

## Leading Questions for Developing Learning Goals

### Questions for Faculty

- In general, what are the most important things a student gains from your field of study?
- What qualities and capabilities do you strive to foster in your students?
- What is the most important knowledge that your students acquire from your field of study or from working with you?
- How does your field of study or your work change the way students view themselves?
- In what ways does your field of study or what you do contribute to a student's well being?
- How does your field or what you do change the way a student looks at the world?
- What does your field of study or what you do contribute to the well being of society at large?
- How do people in this area of study differ from those in other areas (knowledge, skills, and/or values)?
- How do we know the extent to which students are learning what we hope from our field of study?
- How do we use information about student learning and development to enhance student learning?

### Questions for Students

- What is the most important knowledge you have gained from taking courses, minoring, or majoring in this subject?
- What are the most valuable skills or abilities that have you developed as a result of taking courses, minoring, or majoring in this subject?
- How has taking courses, minoring, or majoring in this subject changed the way you look at yourself?
- How has taking courses, minoring, or majoring in this subject changed the way you look at the world?
- How has taking courses, minoring, or majoring in this subject changed the way you think about the future?
- How do you know whether these changes have occurred?
- How do people in this area of study differ from those in other areas (knowledge, skills, and/or values)?
- What changes might be made in course and programs of your major or minor to enhance student learning?

Based on leading questions developed by Prof. C. Ewart, Department of Psychology, Syracuse University, 1998. Reproduced with permission.

Figure 4

## Teaching Goals Inventory Self-Scorable Version

### Purpose

The *Teaching Goals Inventory* (TGI) is a self-assessment of instructional goals.

Its purpose is three-fold: (1) To help college teachers become more aware of what they want to accomplish in individual courses; (2) To help faculty locate Classroom Assessment Techniques they can adapt and use to assess how well they are achieving their teaching and learning goals; and, (3) To provide a starting point for discussions of teaching and learning goals among colleagues.

### Directions

Please select ONE course you are currently teaching. Respond to each item on the Inventory in relation to that particular course. (Your responses might be quite different if you were asked about your *overall* teaching and learning goals, for example, or the appropriate instructional goals for your discipline.)

Just to remind yourself, please print the title of the specific course you are focusing on below:

---

Please rate the importance of each of the 52 goals listed below to the specific course you have selected. Assess each goal in terms of what you deliberately aim to have your students accomplish, rather than in terms of the goal's general worthiness or overall importance to your institution's mission. There are no "right" or "wrong" answers; only personally accurate or inaccurate ones.

For each goal, circle *only one* response on the 1 to 5 rating scale. You may find it helpful to quickly read through all 52 goals *before* rating their relative importance.

In relation to the course you are focusing on, indicate whether each goal rated is:

- |                           |   |
|---------------------------|---|
| (5) <b>Essential</b>      | A goal you <i>always/nearly always</i> try to achieve (76% to 100% of the time) |
| (4) <b>Very Important</b> | A goal you <i>very often</i> try to achieve (51% to 75% of the time)            |
| (3) <b>Important</b>      | A goal you <i>sometimes</i> try to achieve (26% to 50% of the time)             |
| (2) <b>Unimportant</b>    | A goal you <i>rarely</i> try to achieve (1% to 25% of the time) or              |
| (1) <b>Not Applicable</b> | A goal you <i>never</i> try to achieve.   |

Please note: This Inventory was developed with support from The Pew Charitable Trusts and the Ford Foundation by K. P. Cross & T. A. Angelo, U. C. Berkeley School of Education, 1992. Reproduced with permission of the authors.

Rate the importance of each goal below in terms of what you aim to have students accomplish in your course.	<i>Essential</i>	<i>Very Important</i>	<i>Important</i>	<i>Unimportant</i>	<i>Not Applicable</i>
1. Develop ability to apply principles and generalizations already learned to new problems and situations	5	4	3	2	1
2. Develop analytic skills	5	4	3	2	1
3. Develop problem-solving skills	5	4	3	2	1
4. Develop ability to draw reasonable inferences from observations	5	4	3	2	1
5. Develop ability to synthesize and integrate information and ideas	5	4	3	2	1
6. Develop ability to think holistically: to see the whole as well as the parts	5	4	3	2	1
7. Develop ability to think creatively	5	4	3	2	1
8. Develop ability to distinguish between fact and opinion	5	4	3	2	1
9. Improve skill at paying attention	5	4	3	2	1
10. Develop ability to concentrate	5	4	3	2	1
11. Improve memory skills	5	4	3	2	1
12. Improve listening skills	5	4	3	2	1
13. Improve speaking skills	5	4	3	2	1
14. Improve reading skills	5	4	3	2	1
15. Improve writing skills	5	4	3	2	1
16. Develop appropriate study skills, strategies, and habits	5	4	3	2	1
17. Improve mathematical skills	5	4	3	2	1
18. Learn terms and facts of this subject	5	4	3	2	1
19. Learn concepts and theories in this subject	5	4	3	2	1
20. Develop skill in using materials, tools, and/or technology central to this subject	5	4	3	2	1
21. Learn to understand perspectives and values of this subject	5	4	3	2	1
22. Prepare for transfer or graduate study	5	4	3	2	1

K. P. Cross & T. A. Angelo, U.C. Berkeley School of Education, 1992

Rate the importance of each goal below in terms of what you aim to have students accomplish in your course.	<i>Essential</i>	<i>Very Important</i>	<i>Important</i>	<i>Unimportant</i>	<i>Not Applicable</i>
23. Learn techniques and methods used to gain new knowledge in this subject	5	4	3	2	1
24. Learn to evaluate methods and materials in this subject	5	4	3	2	1
25. Learn to appreciate important contributions to this subject	5	4	3	2	1
26. Develop an appreciation of the liberal arts and sciences	5	4	3	2	1
27. Develop an openness to new ideas	5	4	3	2	1
28. Develop an informed concern about contemporary social issues	5	4	3	2	1
29. Develop a commitment to exercise the rights and responsibilities of citizenship	5	4	3	2	1
30. Develop a lifelong love of learning	5	4	3	2	1
31. Develop aesthetic appreciations	5	4	3	2	1
32. Develop an informed historical perspective	5	4	3	2	1
33. Develop an informed understanding of the role of science and technology	5	4	3	2	1
34. Develop an informed appreciation of other cultures	5	4	3	2	1
35. Develop capacity to make informed ethical choices	5	4	3	2	1
36. Develop ability to work productively with others	5	4	3	2	1
37. Develop management skills	5	4	3	2	1
38. Develop leadership skills	5	4	3	2	1
39. Develop a commitment to accurate work	5	4	3	2	1
40. Improve ability to follow directions, instructions, and plans	5	4	3	2	1
41. Improve ability to organize and use time effectively	5	4	3	2	1
42. Develop a commitment to personal achievement	5	4	3	2	1
43. Develop ability to perform skillfully	5	4	3	2	1

K. P. Cross & T. A. Angelo, U.C. Berkeley School of Education, 1992



Rate the importance of each goal below in terms of what you aim to have students accomplish in your course.	<i>Essential</i>	<i>Very Important</i>	<i>Important</i>	<i>Unimportant</i>	<i>Not Applicable</i>
44. Cultivate a sense of responsibility for one's own behavior	5	4	3	2	1
45. Improve self-esteem/self-confidence	5	4	3	2	1
46. Develop a commitment to one's own values	5	4	3	2	1
47. Develop respect for others	5	4	3	2	1
48. Cultivate emotional health and well-being	5	4	3	2	1
49. Cultivate physical health and well-being	5	4	3	2	1
50. Cultivate an active commitment to honesty	5	4	3	2	1
51. Develop capacity to think for one's self	5	4	3	2	1
52. Develop capacity to make wise decisions	5	4	3	2	1

### Self-Scoring Worksheet

1. In all, how many of the 52 goals did you rate as "Essential"? \_\_\_\_\_
2. How many "Essential" goals did you identify in each of the six clusters listed below?

Cluster Number and Name	Goals included in cluster	Total number of "Essential" goals in each cluster	Clusters Ranked (1st to 6th) by number of "Essential" goals included
I. Higher-Order Thinking Skills	1 - 8	_____	_____
II. Basic Academic Success Skills	9 - 17	_____	_____
III. Discipline-Specific Knowledge & Skills	18-25	_____	_____
IV. Liberal Arts & Academic Values	26-35	_____	_____
V. Work and Career Preparation	36-43	_____	_____
VI. Personal Development	44-52	_____	_____

# 3

## Evaluating Student Learning

**T**here are many ways to approach the evaluation of student learning. The characteristics of good evidence of student learning include considerations of direct and indirect methods for gathering evidence of student learning, the appropriate use of quantitative and qualitative evidence, and other methodological considerations. First, however, it is important to understand the fundamental assessment concepts of formative and summative assessment and benchmarking.

### *Formative and Summative Assessment*

Formative assessment is ongoing assessment that is intended to improve an individual student's performance, student learning outcomes at the course or program level, or overall institutional effectiveness. By its nature, formative assessment is used internally, primarily by those responsible for teaching a course or developing a program.

Ideally, formative assessment allows a professor, professional staff member, or program director to act quickly to adjust the contents or approach of a course or program. For instance, a faculty member might revise his or her next unit after reviewing students' performance on an examination at the end of the first unit, rather than simply forging ahead with the pre-designated contents of the course. Many instructors also solicit repeated brief evaluations of their teaching, and the data gleaned from these can be used to make adjustments that may improve learning, such as the introduction of more discussion into a class.

In contrast, summative assessment occurs at the end of a unit, course, or program. The purposes of this type of assessment are to determine whether or not overall goals have been achieved and to provide information on performance for an individual student or statistics about a course or program for internal or external accountability purposes. Grades are the most common form of summative assessment.

Goals for student learning will be expressed summatively when faculty members are describing what they expect students to be able to do or what skills they expect students to achieve when they complete a course or a program or when they graduate from the institution.

Formative and summative assessment work together to improve learning. They should be central components of assessment at the course level, and where appropriate, at the program level.

### *Benchmarking*

The term benchmarking is now common in assessment plans and conversations about assessment. Originally, benchmarking was a term used in the corporate environment to define a set of external standards against which an organization could measure itself. The organization identifies comparable, peer, or "reach" organizations and systematically compares its practices or achievements against those of the other organization.

In higher education settings, a university might use benchmarking techniques to define its comparison group—its peer institutions—and to compare its

own outcomes to theirs. This benchmarking could be based, for example, on retention rates, five-year graduation rates, admissions yield data (the number of enrollees as a function of the number of students accepted), employment and graduate school placement rates, and performance on national or professional examinations. Theoretically, any outcome for which there are data from peer institutions and programs can be compared in a benchmarking study.

Two other related forms of benchmarking are used in higher education settings. A college or university can compare itself to a national norm by reviewing the data from a published test or survey such as the National Survey of Student Engagement (NSSE). Alternatively or in addition, an institution can set for itself the goals or benchmarks that it hopes to achieve within a specified time period (e.g., to increase job placement rates from 70% to 90% in five years).

The benefit of inter-institutional comparison is that it can flag problem areas to investigate the causes of results that differ from the norm. For example, two comparable liberal arts colleges with similar selectivity, similar student preparedness, similar socioeconomic profiles for their students, and similar science curricula, may discover that proportionately more students are accepted to medical schools from one institution than from another. Further investigation may reveal that the excelling college requires a hospital internship for all of its pre-med students.

The discovery that an institution's students are below the norm on a national metric (e.g., amount of time devoted to school work outside the classroom) challenges the institution to determine the reason for this result. Similarly, an institution that sets its own internal benchmarks must design and implement a program to achieve its goals.

Before beginning to articulate goals for student learning, program faculty and leaders of institutional assessment should consider how the use of benchmarks could enhance their institution's ability to achieve its goals and whether useful measures from comparable peer institutions are available.

## ***Direct and Indirect Methods for Assessing Student Learning***

The concepts of direct and indirect methods of evaluating student learning are often confused with each other and with quantitative and qualitative forms of information. Each of these has its merits and drawbacks.

Direct and indirect methods of evaluating learning relate to whether or not the method provides evidence in the form of student products or performances. Such evidence demonstrates that *actual learning* has occurred relating to a specific content or skill. Indirect methods reveal characteristics associated with learning, but they only imply that learning has occurred. These characteristics may relate to the student, such as perceptions of student learning, or they may relate to the institution, such as graduation rates.

When a student completes a calculus problem correctly and shows her work, learning is demonstrated *directly*. When the same student describes her own calculus abilities as excellent, she is demonstrating *indirectly* that she has learned calculus. Both of these pieces of information about the student's performance are important. For example, a student's perception that she is doing poorly in calculus when she is actually doing well would provide important information to both the student and the professor. However, indirect evidence—in this case, a perception—is less meaningful without the associated direct and tangible evidence of learning.

Figure 5 includes examples of direct and indirect measures of student learning at the course, program, and institutional levels. Many of the examples presented in Figure 5 can be used as measures of student learning at more than one level. For example, portfolios of student work and student satisfaction surveys can be used at the course, program, and institutional level, and internship performance ratings could be used at the course or program level.

Figure 5

## Examples of Direct and Indirect Measures of Student Learning (Course, Program, and Institutional Levels)

	Direct Measures	Indirect Measures
<b>Course</b>	<ul style="list-style-type: none"> <li>* Course and homework assignments</li> <li>* Examinations and quizzes</li> <li>* Standardized tests</li> <li>* Term papers and reports</li> <li>* Observations of field work, internship performance, service learning, or clinical experiences</li> <li>* Research projects</li> <li>* Class discussion participation</li> <li>* Case study analysis</li> <li>* Rubric (a criterion-based rating scale) scores for writing, oral presentations, and performances</li> <li>* Artistic performances and products</li> <li>* Grades that are based on explicit criteria related to clear learning goals</li> </ul>	<ul style="list-style-type: none"> <li>* Course evaluations</li> <li>* Test blueprints (outlines of the concepts and skills covered on tests)</li> <li>* Percent of class time spent in active learning</li> <li>* Number of student hours spent on service learning</li> <li>* Number of student hours spent on homework</li> <li>* Number of student hours spent at intellectual or cultural activities related to the course</li> <li>* Grades that are not based on explicit criteria related to clear learning goals</li> </ul>
<b>Program</b>	<ul style="list-style-type: none"> <li>* Capstone projects, senior theses, exhibits, or performances</li> <li>* Pass rates or scores on licensure, certification, or subject area tests</li> <li>* Student publications or conference presentations</li> <li>* Employer and internship supervisor ratings of students' performance</li> </ul>	<ul style="list-style-type: none"> <li>* Focus group interviews with students, faculty members, or employers</li> <li>* Registration or course enrollment information</li> <li>* Department or program review data</li> <li>* Job placement</li> <li>* Employer or alumni surveys</li> <li>* Student perception surveys</li> <li>* Proportion of upper-level courses compared to the same program at other institutions</li> <li>* Graduate school placement rates</li> </ul>
<b>Institutional</b>	<ul style="list-style-type: none"> <li>* Performance on tests of writing, critical thinking, or general knowledge</li> <li>* Rubric (criterion-based rating scale) scores for class assignments in General Education, interdisciplinary core courses, or other courses required of all students</li> <li>* Performance on achievement tests</li> <li>* Explicit self-reflections on what students have learned related to institutional programs such as service learning (e.g., asking students to name the three most important things they have learned in a program).</li> </ul>	<ul style="list-style-type: none"> <li>* Locally-developed, commercial, or national surveys of student perceptions or self-report of activities (e.g., National Survey of Student Engagement)</li> <li>* Transcript studies that examine patterns and trends of course selection and grading</li> <li>* Annual reports including institutional benchmarks, such as graduation and retention rates, grade point averages of graduates, etc.</li> </ul>

## Direct Methods

Direct methods of evaluating student learning are those that provide evidence of whether or not a student has command of a specific subject or content area, can perform a certain task, exhibits a particular skill, demonstrates a certain quality in his or her work (e.g., creativity, analysis, synthesis, or objectivity), or holds a particular value. Direct methods can be used at the course level, the program level, and, theoretically, at the institutional level.

**Course Level.** Most familiar are direct evaluations of learning that are applied at the course level. Examinations,<sup>3</sup> regardless of format, are designed to be direct evaluations of student learning. Similarly, evaluations of writing samples, presentations, artistic performances, and exhibits provide direct evidence of student learning, as do evaluations of student performance in internships, research projects, field work, or service learning settings. As discussed later, grading linked to clear learning goals is a valid and useful form of direct measurement of student learning.

**Program Level.** At the program level, examinations also are used frequently as direct measures of student learning. Such examinations would be more comprehensive than those embedded within a course and would be designed to evaluate cumulative, aggregate, or holistic learning after the conclusion of a program or during the course of the program.

For example, a writing examination might be given after the first two years of a general education program, with the goal of determining whether students' writing was enhanced as a function of the program. Standardized tests of disciplinary content might be administered to students after they have completed all program requirements for the major (e.g., American Chemical Society examinations). Honors theses, senior theses, or senior projects are other sources of direct evidence of student learning within a program. Ratings by internship supervisors

of how well interns are demonstrating key learning outcomes are important, direct program-level evidence of student learning.

**Institutional Level.** Direct evaluations at the institutional level are used less frequently and are much more likely to take the form of an examination. A college or university might use the Academic Profile or the Major Field Tests from the Educational Testing Service, the Collegiate Assessment of Academic Proficiency from the ACT (American College Testing) or other graduate-level examination scores to demonstrate that learning has occurred.

An institution may wish to demonstrate that certain goals expressed in its mission were achieved through exposure to the entirety of its curriculum and co-curricular experiences. For example, it may wish to show that regardless of program or major, which co-curricular activities students have participated in, and whether students were residents or commuters, they exhibit cultural sensitivity and global cultural and geographical awareness. It could design an evaluation process to determine the degree to which graduating students exhibited these qualities (e.g., a rubric for reviewing an examination or portfolio).

It may appear that such qualities are abstract and, therefore, that the measurement of learning was not direct, but in fact that is not the case. In this example, the goal was to have students learn, through curricular and co-curricular programs, to be good global citizens, broadly speaking, and the hypothetical examination was designed to measure the degree to which this goal was achieved.

General education knowledge, competencies, and skills gained across the curriculum might be evaluated over the entire student experience, whether before or after graduation.

---

3 For the purpose of clarity, the term "examination" is being used here to refer to what are commonly called quizzes, exams, or tests designed to measure whether or not a student has learned something that he or she was taught prior to its administration. The word "test" is a more generic term and can apply to any measure that may be direct or indirect, or qualitative or quantitative.

**Fundamental Importance of Direct Forms of Evaluation.** The power of direct assessments of student learning is that, if designed properly, they answer the most important questions:

- What did students learn as a result of an educational experience?
- To what degree did students learn?
- What did students *not* learn?

Institutional stakeholders and the public can understand easily data gleaned from direct evaluations of student learning. They can understand, for instance, that students at Institution A have higher scores on the American Chemical Society examination than students at Institution B, and those same data provide assurance that a certain level of knowledge has been acquired by students at both institutions.

**Limitations and Considerations Related to Direct Forms of Evaluation.** Direct assessments, however, do not tell the whole story of student learning. There are two potential problems with using only direct assessments of student learning. The first problem relates only to direct methods, and the second pertains to both direct and indirect methods.

Direct assessments of student learning, while providing evidence of *what* the student has learned, provide no evidence as to why the student has learned or *why* he or she has not learned. The “why” of student learning is especially important when students have not learned, because one of the primary goals of assessment is to make future learning experiences more effective.

If students perform poorly on a mathematics exam, for instance, it is important for the instructor to know whether the students’ performance resulted from not having learned the material or from having learned the material but also experiencing anxiety during the examination. Other data are needed to answer this question.

It is important to consider that even direct forms of evaluation do not necessarily indicate whether “value-added” learning has occurred. The Commission does not require that its member institutions demonstrate value-added learning, only

that the institution’s learning outcomes are consistent with its goals.

In and of themselves, direct forms of evaluation do not always provide evidence that the targeted learning goal was achieved within the context of a course, a program, or an entire college education, or whether the demonstration of the learning goal was influenced by or a product of prior learning or even the result of innate abilities. If an institution or faculty members in a program are concerned about demonstrating that the learning occurred in a particular context, then care should be taken to design aspects of the assessment program to tap “value-added” learning.

At the course level, the contrast between value-added demonstrations of student learning and absolute levels of student learning is rarely meaningful. One can assume, for instance, that knowledge of college-level organic chemistry, elementary school teaching techniques, or Spinoza’s philosophy was acquired within the context of the course specifically designed to teach that knowledge. The same reasoning applies to the program level; students are likely to have acquired the skills and knowledge specific to their programs while taking courses within the program.

At the institutional level, the distinction between student knowledge that was acquired before the student arrived at the institution and what he or she learned while in attending the institution may be a more salient one. Some institutions may want to demonstrate that the education they provide has had a fundamental effect on students’ lives—i.e., changed them in a way that would not have occurred if the student did not attend college or attended a different type of college.

One college, for instance, may want to demonstrate that a personal atmosphere that encourages faculty-student mentoring relationships results in better preparation for acceptance to graduate school than a student might otherwise receive at a different type of institution. Another may want to demonstrate that it prepares its students for the real world in a way that a different college experience cannot. Yet another might use assessment data to show that students have dramatically increased their job marketability or their chances of graduating by attending the college.

If institutions seek to demonstrate such accomplishments, it is important to consider whether the assessment design truly demonstrates value-added learning rather than some other phenomenon. For instance, students entering an institution with very high SAT writing scores are likely to write well after they have been exposed to the college's General Education program. In other words, to the extent that high scores of graduating students on tests of writing skills reflect pre-college expertise, those scores reflect the effect(s) of one or more "inputs" but are not necessarily value-added.

Value-added gains can be useful in assuring the college community and the public that higher education provides cognitive, affective, and social growth beyond the levels the students had attained when they entered college. However, devoting too much attention to creating an assessment design that rules out other causes for student learning can take the focus away from the most important question: Have students who graduate from this college or university learned what the institution hoped they would learn?

### **Indirect Methods**

Indirect methods of evaluating student learning involve data that are *related to* the act of learning, such as factors that predict or mediate learning or perceptions about learning but do not reflect learning itself. Indirect evidence often is acquired through the use of self-report format surveys, questionnaires, and interviews. Indirect evidence also is provided in the form of "demographic" statistics about the student population of an institution, such as overall GPA, student retention rates, graduation rates, and job placement rates. Qualitative information about graduates, such as names of employers, graduate schools attended, and alumni achievements are also common forms of indirect evidence.

**Course Level.** The most familiar indirect assessment of student learning is the course and/or teaching evaluation given at the end of the semester. These instruments usually have a quantitative section in a Likert (numerically-scaled) format, in which the student rates the quality of teaching and of the course, as well as a narrative section in which the student offers additional qualitative comments.

An instructor who regularly reviews his or her teaching evaluations and who changes the course as a result of those evaluations is engaging in improvement based on hypotheses derived from the indirect assessment of student learning. The same instructor can use this indirect method in conjunction with direct methods to improve student learning in the course.

For example, students might use the narrative portion of the evaluation to request more time for class discussion and might give the professor only moderate ratings for "engagement with the course material." The instructor decides to introduce more discussion into his or her class and subsequently students praise the use of discussion and give high ratings for the instructor's "engagement with course material." Most importantly, the instructor notices that student grades on quizzes or exams and work on assignments are higher in the semester after he made the change. This simple illustration of how indirect evidence can be used in conjunction with direct evidence can be applied in more complicated situations.

**Program Level.** At the program level, student satisfaction surveys may reveal that students want more one-on-one contact with faculty members. Upon reflection, faculty members may decide to offer more independent study experiences; consequently, scores on Graduate Record Examination subject area exams improve (direct evidence of student learning), as do graduate school admission rates (indirect evidence of student learning).

**Institutional Level.** Indirect means of evaluating student learning are important at the institutional level as well. National surveys, such as the National Survey of Student Engagement (NSSE), provide benchmarking opportunities for the institutions to gauge the qualities of their student populations relative to other institutions so that they can determine whether changes in programming affect students' perceptions and behavior inside and outside the classroom. Ultimately, such assessments can affect performance in the classroom.

For instance, if an institution finds that its students spend less time studying than the national average for study time, it might introduce curricular changes that link student evaluation (i.e., grades)

more directly to the amount of time studied, perhaps by providing assignments that demand more out-of-class time and by using class examinations which test areas that are not learned simply by attending class. The greater engagement that these changes create might serve to improve student performance on direct measures of student learning.

Indirect evidence often focuses on the learning *process* and the learning *environment*. Alexander Astin's input-environment-output assessment model, based on research from the past several decades (e.g., Astin, 1991; Chickering & Gamson, 1991; Pascarella & Terenzini, 1991) indicates that students learn most effectively when, in general, they are engaged in the learning process and they can see a connection among course goals, course content, and evaluation.<sup>4</sup>

The extent to which these inputs and processes exist may support the inference that student learning is taking place. Each of these discoveries about student learning was gained through indirect methods of assessment, such as surveys of student perceptions and opinions. The results of these surveys then were correlated with actual student learning outcomes (measured directly), demonstrating that when the existence of specified inputs and processes correlates with student learning.

**Limitations and Considerations Related to Indirect Methods of Evaluation.** The most important limitation of indirect methods is that they do not evaluate student learning *per se*, and therefore should not be the only means of assessing outcomes.

As with direct measures of student learning, it is important to consider that indirect measures do not necessarily imply that value-added learning has occurred. Students who express indifference to co-curricular activities after their first year may be expressing an indifference that is the result of dissatisfaction with campus programs, or they may have arrived on campus disinclined to spend time on anything but course-related work.

As noted above, the Commission does not require proof of value-added student learning.

Nevertheless, an institution should consider whether value-added data are necessary to demonstrate that it fulfills its own mission. If so, it should ensure that data collection procedures warrant conclusions about the effectiveness of programs in teaching students.

### ***Quantitative vs. Qualitative Evidence***

In every example of direct and indirect assessment cited above, the evaluation of student learning could provide either qualitative or quantitative information. Both qualitative and quantitative information are valuable forms of evidence about student outcomes.

Quantitative evidence consists of data that are represented numerically. For instance, performance on a test or responses to a questionnaire may be scored so that a number represents the degree to which an individual performed or agreed/disagreed with a certain concept. Because quantitative data are expressed in numbers, they can be compared directly or subjected to statistical analysis, and they can enable the researcher to make certain assumptions when comparing one data point to another. Quantitative data also may permit one to express numerically meaningful changes in performance (given certain conditions). One may claim, for instance, that a change in a test score from 50 to 60 represents a 10-point or a 20 percent gain in an individual's performance, expressed as a percentage of his or her original score. Quantitative data, therefore, are valued for the ease with which calculations and comparisons can be made, and for the easily understandable representations of performance that they produce.

Qualitative evidence typically comes in two forms. The first form involves simple categorization of individuals into discrete groups (e.g., employed or unemployed; participates in athletics or does not participate in athletics). The second form of qualitative evidence is data expressed in prose or narrative. A question is asked of an individual and

---

<sup>4</sup> See Chapter 5 for a further discussion of this topic. See also Figure 21 for a list of conditions under which students learn most effectively.



he or she responds in a free-form manner, expressing, for instance, an idea, opinion, or evaluation. Because of their non-numerical nature, qualitative data cannot be subjected directly to statistical analyses, nor can easy direct comparisons be made without engaging in an intervening process to categorize or interpret the data. Qualitative data, however, can be “richer” than quantitative data, because they provide a more extensive variety of information related to a particular learning goal. Many faculty members, for instance, use the numerical scores (quantitative data) from their teaching evaluations to make overall judgments of their own performance, but they value the qualitative, narrative comments from students as more useful in revealing students’ personal perceptions of a course.

A common misconception is that qualitative assessments are not as reliable, valid, or objective as quantitative ones. This is not necessarily the case. There are well-designed and statistically reliable means of interpreting and analyzing qualitative data and numerous resources for learning to use qualitative methods (see Silverman, 2001; Maxwell, 1996). For example, an instructor might assess the same learning goals using a multiple-choice test or an essay test. Similarly, an instructor might grade a senior project presentation quantitatively with a standard set of evaluation criteria (i.e., a rubric). Alternatively, he or she might provide the student with a prose evaluation, in a non-scaled format, citing the strengths and weaknesses of the presentation. However, it is best if this evaluation is organized around a standard set of criteria that were shared with the student beforehand.

A student survey designed to gather information on student satisfaction may elicit data that are quantitative (i.e., “On a scale of 1 to 7, how satisfied are you with the quality of advising?”) or qualitative (“How would you describe your experience with academic advising?”). Similarly, employers asked to assess the strengths and weaknesses of alumni may be asked to assign “scores” to, or to describe, alumni characteristics.

Most beginning assessment initiatives are likely to rely more heavily on quantitative, rather than qualitative, forms of assessment for several reasons. Quantitative data are easier to collect and are in the

form of a readily-analyzable numeric score. In contrast, qualitative data must be sorted, categorized, and interpreted (most often by humans rather than by computer programs) before a final judgment can occur. Methods of ensuring the reliability of qualitative data are time-consuming. For instance, to ensure that portfolio assessment is reliable, at least two raters are used to review each portfolio, providing a form of “inter-rater” reliability. Focus groups, another commonly used form of qualitative data collection, require large investments of time to gather data from comparatively few students.

A good use of qualitative evaluation is to help develop quantitative evaluation criteria (rubrics). For instance, one might conduct focus groups for the purpose of designing questions for a satisfaction questionnaire or use a scoring rubric for portfolios to determine what characteristics of students’ writing might be evaluated.

For assessing student learning, *Characteristics* encourages the use of multiple approaches—both quantitative and qualitative—but it does not *require* the use of both approaches (see Standard 14). Institutions and faculty members in different programs should be thoughtful about which approach, or combination of approaches, best suits the student outcomes that are being assessed in each unique situation.

### ***Other Methodological Considerations***

Some of the other methodological considerations often raised with regard to assessment include reliability and validity; pretests, posttests, and longitudinal design; the role of grades, self-report measures, and statistical versus practical significance.

### **Validity and Reliability**

In general, the terms “validity” and “reliability” refer to the extent to which assessment tools and methods provide accurate, fair, and useful information. Both concepts are important factors in choosing standardized assessment instruments and should be considered seriously when developing locally-created instruments for summative assessment.

Validity refers to the integrity of the instrument. Does the instrument measure what it was designed to measure, or does it actually measure something else? An instrument designed to assess student sensitivity to the cultural norms of others, for instance, may actually be measuring a student's sensitivity to detecting those responses desired by the professor or the institution that values such sensitivity. Obviously, the instrument would not provide a valid assessment of cultural sensitivity.

Three forms of validity are especially relevant to assessing student outcomes. An instrument with "construct validity" adequately taps the "construct" or conceptual framework that it is designed to measure because its questions have been developed specifically for that purpose. The test of cultural sensitivity described above lacks construct validity because it assesses student perceptions of faculty beliefs, not cultural sensitivity.

Content validity, and in particular "face validity," refers to the content and structure of an evaluation instrument: On the face of it, does it appear to assess what it is designed to assess (Gall, Borg & Gall, 1998). The cultural sensitivity instrument described above may appear to have face validity—the questions appear to be about cultural sensitivity—even though it lacks construct validity. In general, however, the content and structure of an instrument should make sense to those who are using it. Several methods are employed by test designers to ensure that instruments have both content and face validity.

A third important form of validity is referred to as "concurrent" or "criterion validity." Criterion validity means that a test or assessment instrument will yield results that are similar to those of other instruments designed to assess the same outcome. Two tests of college mathematics ability, for instance, should yield similar results when administered to the same students; if one measure of student satisfaction demonstrates that students are very satisfied, another should as well. This result also could demonstrate "predictive validity" if the strength of the correlation between the two measures was great. A test or other evaluation instrument with good criterion validity also will predict performance on other measures of constructs that should be related. For instance, student satisfaction should predict retention, and

high scores on a test of ethical decision-making should predict ethical behavior. Additional concepts and examples related to reliability and validity are discussed in the section below entitled, "Key questions for choosing assessment instruments."

Reliability refers to the consistency of results for a test or assessment instrument over repeated administrations to the same individuals. For instance, an aptitude test for mechanical engineering, given twice to the same person, should yield similar results each time. Otherwise, it fails in its purpose of providing an accurate prediction of future success in mechanical engineering.

Reliability is established during the development of the test, when special populations are recruited to take the test more than once, before the test is used for its intended purpose. Reliability information about standardized tests is presented in the form of statistical correlations (which should be very high) among repeated administrations of the test in the same population.

The concepts of validity and reliability apply primarily to summative assessment, and not as directly to formative assessment, because instructor-created examinations and measures usually only exhibit "face validity," not the other forms of validity discussed here, and they are not usually subjected to rigorous pre-administration tests of reliability.

### **Pretests, Posttests, and Longitudinal Designs**

A common misconception is that, in order to make any claims about "value-added" changes in student learning, one must use a pretest-posttest format. For instance, in order to demonstrate that a general education program has improved the writing skills of students, it appears that it would be necessary to have data on the writing skills of the *same* students before they began the program. This notion could thwart attempts to assess writing skills, and in a large institutional setting, it could be so daunting that it could short-circuit any attempt to demonstrate that writing skills have improved.

Two conceptual alternatives to a pretest-posttest are discussed briefly below. Research methods experts on most campuses could further explain these and suggest additional alternatives.

The first option would be to identify which general education experiences were designed specifically to enhance writing skill. Perhaps these experiences include courses in introductory composition, rhetoric, and an initial writing-intensive course in the major. One then could compare two populations of first-year students or two populations of sophomores—those who had completed these courses with those who had not. The group that has not completed the courses can serve as the comparison or “control” against the group that completed the courses.

A second option is to compare students against a national norm on a standardized test or against a targeted “benchmark” population. Suppose the learning goal in question is that students have gained a certain level of mathematical proficiency as a consequence of taking a two-course mathematics sequence in a general education program. One can administer a standardized test of college-level mathematics after students have completed the sequence and compare students’ scores to national norms. In this case, no pretest was necessary; the national norm serves as the comparison or “control” group. This method is problematic if students at the institution are not drawn from an average population, as would be the case if the institution is highly selective or open-access. However, it does produce meaningful comparisons if an institution’s student population roughly approximates an average population. Scholastic Achievement Test scores, for instance, might be used as a measure of the level of selectiveness used in admitting students.

If the institution’s population is not average, a benchmarking strategy would be a more appropriate alternative. Students’ scores on a test of college mathematics could be compared to the scores of students at institutions with comparable populations. Scores higher than those of the benchmarked school would be convincing evidence that the math curriculum of the target institution is successful.

A common assertion is that longitudinal research designs (those that follow the same individuals over time) are necessary to draw meaningful conclusions about what students have learned. Sometimes a longitudinal perspective is warranted because other approaches are less valid. For example, if an

institution is interested in demonstrating that its graduates are successful in their careers, a longitudinal survey administered to graduates repeatedly over several years would be appropriate for several reasons. Demographic data tell us, for instance, that people change careers multiple times during their lives, so examination of a single “window” of time may not be an accurate assessment. In addition, the population of graduates offers the benefit that its members will be around long enough to be surveyed repeatedly over time.

Most importantly, however, a longitudinal design guards against “cohort effects” that could intrude if graduates from one generation were compared with graduates from another generation. Career trajectories may change historically, and the character of the institution may have been markedly different in the past. Thus, 1950s female graduates may hold a lower percentage of professional degrees than 1980s female graduates. This finding tells us more about historical context than institutional outcomes. However, the same question, asked of the same individuals at several different points in time yields meaningful information. A finding that female students from the same cohort had a greater percentage of graduate degrees 20 years after college than they did 10 years after college could be used (in conjunction with other outcomes data) to demonstrate that the institution produces lifelong learners.

In most cases, when student outcomes during or at the end of a higher education experience are being evaluated, longitudinal data are not necessary and may not yield meaningful information. Pre-test and post-test assessments, as well as alternatives which are discussed above, are more practical alternatives and provide answers to the same general question: “Has meaningful learning occurred as a result of an educational experience?”

### **Where Do Grades Fit into the Picture?**

Faculty members and others often ask whether grades are appropriate and sufficient for assessment of student learning after the learning goals are defined. The answer is that grades have been, and will continue to be, an excellent indicator of student learning *if they are appropriately linked to learning goals*. The Commission recognizes that grades are

an effective measure of student achievement if there is a demonstrable relationship between the goals and objectives for student learning and the particular bases (such as assignments and examinations) upon which student achievement is evaluated (Standard 14).

In and of themselves, however, grades are not direct evidence of student learning. That is, a numeric or a letter grade alone does not express the *content* of what students have learned; it reflects only the degree to which the student is perceived to have learned in a specific context.

One reason “grade inflation” is seen as a problem is that grades alone cannot be relied on to reflect student performance accurately. One could ask: “Does one grade of ‘A’ equal another?” If instructors were to match grades explicitly with goals, it would become easier to combat grade inflation, because high grades must reflect high performance in specified areas.

Grades, however, can provide an excellent means for improving teaching and learning both during a course (formatively) and at the conclusion of a course (summatively). When grades serve as the final judgment of performance in a course, they provide a summative evaluation of students’ performance as individuals and as a class. If the grades of individual students can be traced directly to their respective competencies in a course, the learning achievements of those students are being assessed in a meaningful fashion. If, however, examinations or homework assignments are not designed to test the skills and competencies that the course was designed to teach, then grades for that course are measuring something other than student attainment of the course goals.

Suppose, for instance, an instructor presents the content of an anatomy and physiology course that focuses on identifying and labeling anatomical structures and physiological processes. An appropriate evaluation of student mastery of the course content would be an objective final exam requiring students to label diagrams, answer multiple-choice definitional questions, and fill in the blanks. In contrast, an examination that required students to evaluate a physiology experiment on its methodological merits would not be an assessment of student learning of the course content. Some

students would do well on the essay exam, but their performance probably would not be related to what they learned in the course. In this example, a bad grade could not be attributed to a student’s failure to learn the material or to prepare for the examination. Thus, even the use of grades as a summative assessment warrants a careful approach.

Thoughtfully-constructed syllabi and “test blueprints,” which are discussed later in this chapter, are two of several possible approaches to connecting grades directly to desired course goals.

Grades and grading practices can be a component of formative assessment as well. For example, many faculty members use drafting and revising processes to teach writing. Students mimic the “real world” by writing multiple drafts, submitting them to critiques by the professor or their peers, and revising them for resubmission. Each draft may be assigned a grade in association with critical comments. Depending on the instructor’s preferred strategy, all or only some of the interim grades may be used to determine the final grade. In this case, the grade for each draft, in conjunction with critical comments, gives the student an indication of his or her performance, what might be done to improve the product, and how the quality of the product changes over time.

Grading can be formative when there are multiple means and formats for assessing student learning and when there are repeated opportunities to demonstrate improvement within the context of one class. For instance, a professor might assign two or three papers (with required drafts), two class presentations, two objective format exams, and a journal that must be reviewed repeatedly by the professor during the semester. Each of these “assessment events” could be graded, providing students with at least two distinct types of opportunity to learn more or learn better. A student can compare his or her performance on the various assessment formats, thereby learning which skills he or she has mastered and which should be improved. In addition, a grade on the first test administration or the first paper or presentation serves as feedback (a formative assessment) that provides information on how to improve. This learning experience can be applied toward adapting study skills or work habits before the next attempt.

## Self-report Measures

Concerns are often expressed about the use of self-report measures for answering questions about student learning. Sometimes these concerns relate to the use of indirect methods of assessing student learning and the concerns about qualitative versus quantitative assessment discussed previously. Often, however, concerns are related most directly to the validity and reliability of self-report measures. Self-report measures can be designed to be valid and reliable and can be assessed by applying the characteristics of reliability and validity described above.

Both common sense and face validity should be used to determine the value of a specific self-report measure. For example, if the goal is to determine whether students are satisfied, it seems that a self-report measure is the only means of gathering such data. Satisfaction, by definition, is one's feeling of liking, comfort, and fulfillment resulting from a specific event or situation. Similarly, it is appropriate to gather data on affective states (emotions) and social perceptions with a self-report instrument (assuming that it meets the criteria for reliability and validity).

It is possible to collect direct evidence of student learning using self-report measures, but these must be designed carefully to elicit evidence of student learning. For example, students may be asked to reflect on the most important thing they learned in a specific course, or what else they would like to learn on the same subject. In doing so, they would reveal the actual content of what they had learned. However, self-report questions such as, "Did you learn a lot in this class?" would not elicit such information. Self-report measures are most frequently used to provide valuable indirect evidence of student learning.

## Statistical versus Practical Significance

Data related to student outcomes are often described as being "statistically significant" or "not statistically significant." The concept of statistical significance relates to the probability of a given result occurring by chance. If the result is too unlikely to have occurred by chance, it is said to be statistically significant.

For example, imagine two groups of students, each of whom has completed an introductory calculus course. Assume that members of each group were randomly assigned to two different teaching formats—one problem-based and the other traditional lecture—and that the same professor taught each course. At the completion of the course, both students are given the same standardized calculus examination. The average grade for the students in the problem-based course was 10 points higher than the average grade for the students in the traditional lecture course. Is a 10-point difference enough to make the claim that the problem-based course is a better form of teaching? Would a 2-point difference have been enough? Would 20 points be enough to make the claim? A test of statistical significance would reveal whether the 10-point difference could have happened by accident in a normally distributed population of students (i.e., the difference could have been caused by other factors, unrelated to the course, of which we are unaware), or whether the 10-point difference was large enough that in all likelihood it was caused by differences in the courses.

Judgments of statistical significance only become reliable when there are sufficient numbers of student test or survey results from which to draw inferences about a population of students. In many cases, faculty members will be studying outcomes data from small groups of students or engaging in formative assessment for which ongoing improvement in a class or a program is the goal. In these situations, faculty and staff members should make judgments and introduce changes based on "practical significance." Do the students' scores, or their change in scores from one time to another, reveal a pattern or appear to be meaningful or informative enough to support changes in a course or program?

In general, when large-scale assessments are being used, or when standardized tests are administered program-wide or institution-wide, statistical tests should be used to analyze the data. Guidance may be found in social science, education, mathematics and statistics, and other departments on campus that use empirical methods.

Judgments of student outcomes based on practical significance are equally valid when the number of students being evaluated is small, when data are qualitative rather than quantitative, and when the purpose is to engage in formative assessment.

### ***Key Questions When Choosing and Implementing Evaluation Instruments***

One should ask several questions when choosing assessment instruments:

#### **□ Is the evidence provided by the evaluation method linked to important learning outcomes?**

This is perhaps the single most important way to determine the quality of most evaluation tools and methods. Regardless of whether an evaluation instrument is standardized (previously published and tested for validity and reliability) or “home grown” (created locally for a specific purpose), it is important to ensure that the instrument is designed to provide evidence of the desired learning outcomes. In research design terms, this involves determining whether the operational definition (the aggregate instrument or items on the instrument) actually assesses the construct (the learning goal) that it is intended to assess (construct validity).

For many standardized instruments, the intended purpose will be apparent immediately.

A disciplinary test, for example, such as the American Chemical Society (ACS) test, evaluates students’ knowledge of facts, skills, and procedures that should have been acquired as a function of the undergraduate curriculum in an ACS-accredited program. Subject-area Graduate Record Examinations (e.g., the psychology GRE) evaluate content knowledge in the respective disciplines they represent. Publishers of other standardized tests with other less readily obvious content will explain, in the test information materials, what the test is designed to assess.

It is important, however, not to assume that the linkage between every item on a standardized assessment instrument and the construct it is designed to assess will be readily apparent. Many standardized instruments have built-in reliability checks and so-called “lie-scales.” Measures that

are designed to evaluate affective and social development are especially likely to incorporate a series of questions that seem irrelevant, but that actually enhance the instrument’s validity.

#### **□ Is a standardized instrument appropriate for the learning goals of the institution?**

It certainly is not necessary to use standardized assessment instruments. In fact, for most learning goals, none will be available. Although a test created locally may not have the same statistical validity and reliability as a standardized instrument, its relevance to the specific learning goals in question may make it a more appropriate and effective instrument. A “test blueprint” (an outline that matches test items with the learning outcomes they are intended to assess) can be used to construct a test or instrument or to evaluate how well an existing “home-grown” instrument assesses key learning outcomes.

#### **□ Is the evaluation method appropriately comprehensive?**

No assessment tool or method can assess *every* important learning outcome, but the best ones assess a comprehensive and/or representative sample of key learning outcomes. It is not financially feasible to use several published instruments to assess multiple outcomes, nor is it feasible to subject students to multiple tests or surveys. (The latter has its own measurement problems.) Regardless of whether an assessment instrument is standardized or specially created, it should be as comprehensive as possible.

#### **□ Are important learning outcomes evaluated by multiple means?**

Few evaluation methods are perfect. It is important to triangulate around important learning goals, assessing them through various means, and through tests of various formats. For instance, a standardized test of disciplinary knowledge may be an adequate form of assessment of students’ content knowledge of a discipline, but it may provide no indication of his or her preparedness to be a good practitioner in that discipline.

### **❑ Are the questions clear and interpreted consistently?**

In addition to examining the correspondence between learning goals and the assessment measures being used, it is important to assess whether its “non-content” properties are adequate. For example, a test should not be culture-specific, its vocabulary and sentence structure should be at an appropriate level, and it should not contain ambiguous, unclear, or double-barreled questions (i.e., questions that actually contain two questions).

Questions should be phrased carefully to ensure meaningful responses. For instance, imagine that a targeted learning goal is that students’ desire to engage in community service increases after exposure to a service-learning program. Imagine also the following two questions asked of a graduate:

- ❑ “On a scale of 1 to 7, how likely are you to participate in community service activity?”
- ❑ “On a scale of 1 to 7, how much influence did your community service during college have on your desire to participate in community service in the future?”

Both of these questions are indirect measures of learning or development, but only the second provides information that would help the institution to improve the service-learning program.

A specially created instrument should be reviewed by several colleagues and students to ensure clarity, and it then should be pre-tested by some students who have diverse backgrounds and characteristics in order to clarify ambiguous items.

### **❑ Do questions elicit information that will be useful for making improvements?**

Questions should be designed so that, when possible, they yield responses that both evaluate an aspect of the educational experience and suggest options for improvement. For instance, a survey designed to evaluate student experiences with the Career Services Office should ask about perceptions of its efficacy:

- ❑ “On a scale of 1 to 7, how important was the Career Services Office in helping you find employment upon graduation?”

The instrument also should ask how the office might be improved. For example, the respondent might be asked to name the three most useful activities of the Career Services Office for helping students find jobs and to name three ways in which the functions of that office could be improved.

The concept of creating questions that are useful for making improvements can be applied to direct assessments of student learning as well. For instance, a complicated problem in a physics class can be divided into subsections to help the professor determine which content or skill areas need additional reinforcement.

### **❑ Does everyone interpret the responses the same way?**

When assessments of student outcomes are subjective—that is, if they do not require discrete or quantifiable or unambiguous answers—it is important to develop a rubric (criteria used to score or rate responses) to ensure comparability of review. There should be collegial agreement on what constitutes acceptable, inadequate, and exemplary responses or performance for each assessment instrument to be used, whether it is a paper, a project, a presentation, or an artistic offering. A rubric created to reflect the agreement should be pre-tested by having colleagues independently score the same work samples to see if their scores are consistent. The strategy of inter-rater reliability can be used as well, by enlisting two or more colleagues to “grade” each student’s work or performance.

### **❑ Do the results make sense?**

It is important to use common sense when developing assessment instruments, designing a scoring system or rubric, or interpreting data resulting from assessment instruments. One would expect honors students to outperform other students on their senior theses presentations. One also might expect those same students to fare better in applying to graduate school, but not necessarily in being hired to entry-level positions in corporations. Students who have completed a general education sequence should score better on tests of general knowledge and skills related to specified general education outcomes than students who have not

completed the sequence. Unexpected results should trigger further inquiry.

**□ Are the results corroborated by other evidence?**

It is always important to use multiple means of assessment to determine if a particular learning goal has been met. It also is necessary to compare assessment results for related goals for student learning and even for goals that would be expected to be mutually exclusive. For instance, rubric scores for the writing quality of senior theses should be corroborated by students' grades in composition classes. Faculty ratings and students' self-ratings of performance should correlate with each other. Focus group results should support survey results on the same topic. Conversely, students who demonstrate an increased personal emphasis on wellness by their attendance at the gym and by participation in athletics should not be engaging in increased alcohol and drug consumption. The latter finding would warrant re-evaluation of the campus wellness program.

**□ Are efforts to use “perfect” research tools balanced with timeliness and practicality?**

Although institutions will do their best to ensure that the research designs they use yield meaningful results, they should remember that assessment cannot wait for the perfect research strategy. Indeed, there probably is no perfect strategy. For the purpose of managing the quality and change of an academic curriculum, assessment is a form of systematic inquiry—i.e., “action research” or “applied research,” based on the collection and analysis of data about student learning that is undertaken with the best knowledge and resources permissible and within the time available. The resulting information guides decision makers in choices related to the curriculum, faculty, the use of physical space, and other areas that may have an effect on learning.

**□ Is evidence gathered over time and across situations?**

Assessment is not a once-and-done process. As students, faculty members, curricula, and teaching methods evolve over the years, even institutions with very positive assessment results should undertake repeated assessments to ensure that students are learning as effectively today as they were a few years ago. Because each evaluation technique has relative strengths and weaknesses, there is no single perfect assessment that will yield absolutely accurate information and that is relevant to every situation. In order to have support the findings that each evaluation yields, more than one assessment strategy should be used to corroborate findings.

**□ How much should be assessed?**

Plunging immediately into assessing a large number of students on a full range of learning outcomes will overwhelm faculty members and institutional resources. It will produce an overwhelming amount of information that may be impossible to interpret or to use in enhancing a program. It makes more sense to begin with a more limited approach. For example, faculty members assessing student writing skills might gain more from a thorough analysis of a sample of 30 papers than from a more perfunctory review of 300, as well as by assessing only a few key goals.

Just as every possible outcome need not be measured, it is not necessary to collect data about each student's performance. The Commission is interested in the institution's ability to graduate students with appropriate knowledge, skills, and behavior, not in a demonstration that every student was tested. Meaningful and representative sub-populations (randomly chosen when appropriate) can provide the basis for demonstrating that students across the institution are achieving learning goals.



---

---

## Evaluating Student Learning

- Use indirect measures to explain or support findings from direct measures.
- Choose the most relevant level for evaluation of the learning goals: institution, program, or course.
- Select quantitative or qualitative measures based on type of student learning goals.
- Ensure that grades are related directly to goals.
- Choose appropriate research design.
- Use formative assessment “mid-course” to improve teaching and learning.
- Use common sense: Is the result logical?

---

---

### □ Are faculty and staff members who are knowledgeable about measurement serving as resources for developing assessment instruments?

The work of assessing student learning is essentially systematic inquiry in the tradition of social science or evaluation research, with its attendant need for validity, reliability, control, analysis, and interpretation, to the extent that these are possible. Although everyone involved in the enterprise is an expert in the content base of what is being researched (i.e., teaching and interacting with students in a higher education setting), few are expected to be experts in conducting research. While much of the work of assessing student learning has a common-sense base, it is also true that meaningful analysis of student learning, especially beyond the course level, requires expertise. There are usually faculty members on campus who are trained as social science, education, or other researchers. They can conduct careful, meaningful research and can construct measures. These faculty members, who can be found in psychology, sociology, education, business, and other departments, may be enlisted

to serve as internal consultants, reviewers, statisticians, and mentors in the assessment process.

## *Easy-to-Implement Tools and Techniques*

The assessment tools and techniques presented below yield useful information and are relatively easy to implement. They are not meant to be an exhaustive selection of tools but, rather, an overview of available options.

### Rubrics or Rating Scales

A rubric is an instrument based on a set of criteria for evaluating student work. Rubrics help a professor or other evaluator to make explicit, objective, and consistent the criteria for performance that otherwise would be implicit, subjective, and inconsistent if a single letter grade were used as an indicator of performance. Rubrics delineate what knowledge, content, skills, and behaviors are indicative of various levels of learning or mastery. Ideally, “grading” rubrics are shared with students before an exam, presentation, writing project, or other assessment activity. Conscious awareness of what he or she is expected to learn helps the student organize his or her work, encourages self-reflection about what is being learned and how it is being learned, and allows opportunities for self-assessment during the learning process. Huba and Freed (2000) suggest that instructors consider involving students in the development of rubrics as a class progresses as a way of helping students to develop their own conceptions of what constitutes good and poor work. Both Huba and Freed (2000) and Walvord and Anderson (1998) offer extensive information on developing rubrics.

Figure 6 includes a description of the characteristics and components of rubrics. Huba and Freed (2000) present a thorough description of the uses and purposes for rubrics, along with a comprehensive primer on how to construct them.

There are four basic types of rubrics: simple checklists, simple rating scales, detailed rating scales, and holistic rating scales.

Figure 6

## Criterion-based Rating Scales (Rubrics)

**What is a rubric?** A rubric is a criterion-based rating scale that can be used to evaluate student performance in almost any area. A rubric establishes the “rules” for the assignment (Huba and Freed, 2000). It contains *a priori* criteria for various levels of mastery of an assignment.

**How is a rubric used?** The person evaluating student performance uses a rubric as the basis for judging performance. Ideally, rubrics are available to students prior to their completion of the assignment so that they have clear expectations about the components of the evaluation and what constitutes exemplary performance.

**What are some of the criteria that may be used within a rubric to evaluate student work?** Criteria can include sophistication, organization, grammar and style, competence, accuracy, synthesis, analysis, and expressiveness, among others.

**What are the components of a rubric?** Huba and Freed (2000) describe the following elements of rubrics:

- Levels of mastery (e.g., unacceptable through exemplary)
- Dimensions of quality (see criteria above)
- Organizational groupings (macro categories for criteria)
- Commentaries (the junctures between levels of mastery and dimensions of quality; e.g., a description of the characteristics of an exemplary organization)
- Descriptions of consequences (components of commentaries that relate to real-life settings and situations).

**Where can I see examples of rubrics and learn more?** Walvoord and Anderson (1998) and Huba and Freed (2000) are both excellent sources of information about the characteristics of rubrics and how to develop them. They also provide examples of various forms of rubrics.

**Simple Checklists.** This form of rubric can be used to record whether the relevant or important components of an assignment are addressed in a student’s work. For instance, a rubric might be used to assess whether a laboratory report contained required sections or whether a writing sample contained all of the assigned parts. A checklist of this sort is categorical, that is, it records whether or not a required aspect of an assignment is present, but it does not record quantitative information about the level of competence a student has achieved or the relative skill level he or she has demonstrated.

**Simple Rating Scales.** This form of rubric records the level of student work or categorizes it hierarchically. It is used, for instance, to indicate whether student work is deficient, adequate, or exemplary, or to assign a numerical “code” to

indicate the quality of student work.

In most cases in which a numerical scale is used, it should contain a clear neutral midpoint (i.e., the scale should contain an odd number of rating points). However, survey designers should determine when this might not be appropriate. Occasionally, such scales are intentionally designed without a midpoint in order to force a non-neutral response.

Figure 7, an excerpt from an employee rating scale, is an example of a simple rating scale that does not provide information about the “value” of different points on the scale.

**Detailed Rating Scales.** Detailed rating scales describe explicitly what constitutes deficient, adequate, or exemplary performance on each criterion. Detailed rating scales are especially

Figure 7

## Excerpt from a Simple Rating Scale

### Employer's Final Performance Evaluation of Knowledge, Skills, and Attitudes (KSAs) of: \_\_\_\_\_

Dear Employer:

The College of Business Economics (CBE) understands the need for its graduates to be broadly trained and ready to perform immediately upon entering the job market, both as individuals and in teams. Therefore, its curriculum contains concrete, measurable, and attainable objectives throughout. As a result, each CBE graduate is expected to perform successfully in the following areas of Knowledge, Skills, and Attitudes.

Please rate your intern or employee's performance only on the areas that apply to his/her job.

**The rating scale is: 5=Excellent; 4=Good; 3=Satisfactory; 2=Fair; 1=Poor; N/A=Not Applicable.**

**Excerpt:**

COMMUNICATION: WRITTEN, SPOKEN, GRAPHIC, AND ELECTRONIC	5	4	3	2	1	n/a
1. Write articulate, persuasive, and influential business reports, proposals, and letters						
2. Make articulate, persuasive, and influential individual and team presentations						
3. Develop graphic, spreadsheet, and financial analysis support for position taken						
4. Display presentation skills						
5. Generate appropriate visual aids						
6. Use correct written structure, spelling, grammar, and organization						
7. Articulate another's viewpoint through verbal and non-verbal cue interpretation						
8. Resolve interpersonal and team conflicts						
9. Negotiate effectively						
THINKING: CRITICAL, CREATIVE, AND INTEGRATED	5	4	3	2	1	n/a
10. Use problem-solving techniques						
11. Use adaptable, flexible thinking						
12. Use critical thinking to produce comprehensive, supported, integrated conclusions						
13. Use creative thinking methods to produce ideas						
14. Distinguish fact from opinion, and critical from non-critical information						
15. Develop several workable solutions to a problem						
16. Show common sense						
17. Demonstrate continuous learning (learning to learn)						

Source: College of Business and Economics, Towson University, November 2001. Adapted with permission.

Some of the other characteristics that could be evaluated in the manner shown in Figure 7 include:

- ★Technology
- ★Ethics and Values
- ★Business Disciplinary Content
- ★Leadership, Entrepreneurship,
- ★ Diversity - International and Demographic
- ★ Practical Excellence
- ★ Job Experience and Career Development

useful when several faculty members are scoring student work, because they communicate common performance standards and therefore make the scores more consistent. Detailed rating scales are useful to present to students when an assignment is given or at the beginning of a semester or even a program. They provide students with a clear description of what they are expected to learn and the criteria upon which their learning will be judged.

Figure 8 is an example of a rubric designed as a detailed rating scale.

**Holistic Rating Scales.** Holistic rating scales define deficient, adequate, or exemplary student work as an aggregate, by assigning a single score to a constellation of characteristics that have been fulfilled to a substantial degree, rather than rating each criterion separately. Holistic rating scales often are used when evaluating student work that may vary so widely in form and content that the same criteria may not apply to all. Capstone projects in an art program, for example, might vary so that they cannot all be judged using the same specific criteria. However, a faculty member could create a generic description of what constitutes exemplary work, adequate work, and so on, regardless of the medium or focus of the work.

Figure 9 is an example of a holistic rating scale.

### **Self-reflection**

Asking students to reflect on what and how they have learned—in other words, to engage in metacognition—has several benefits. Student self-assessments give faculty members useful insights into the learning process, help students integrate what they have learned, and provide students with an understanding of the skills and strategies they need to learn most effectively. Classroom assessment techniques suggested by Angelo and Cross (1993) and other similar self-reflection strategies have the added advantage of taking very little faculty or student time. The student often is asked to write simply a phrase or sentence. Examples of self-reflection questions that might be a useful part of an assessment program are provided in Figure 10.

### **Ratings/Comments from Internship or Research Supervisors**

Programs that place students in practica, such as internships, cooperative education, and student teaching assignments, usually require that the on-site supervisor rate the student on essential knowledge, skills, and attitudes. Such scales are relatively simple to construct (see Figure 7.) Because these experiences require students to integrate and use much of what they have learned in a program, these rating scales are evidence of what students have learned during the program. Brief comments from supervisors also provide valuable insights into the overall strengths and weaknesses of a program.

### **Placement Rates**

For professional programs whose goals include preparing students for a particular career, the proportion of graduates who find positions in that career is important indirect evidence of whether students are learning essential knowledge and skills. If a large proportion of graduates from a teacher education program is successful in finding teaching positions, for example, it is likely that those graduates have the knowledge and skills that school administrators consider important for successful teachers. Similarly, if a program aims to prepare students for graduate study or professional programs—pre-medicine and pre-law programs are examples—the proportion of graduates who are admitted into graduate or professional programs is important evidence that students have learned what graduate programs consider important for success in their programs. Note, however, that placement rates alone do not provide insights into exactly *what* students are learning. Therefore, they are usually insufficient evidence of student learning if used alone.

Figure 8

## Example of a Detailed Rating Scale

This scale is adapted from one used to evaluate a “book journal and review” for a cognitive psychology class. For the assignment, students were expected to read one full-length book, chosen from a list provided by the instructor and related to the content of the course but not included on the required course reading list. The purpose of the assignment was to provide a basis for making connections between the course content, other professional or popular work in the field, and students’ daily exposure to topics or situations related to cognitive psychology in their personal lives and in their other courses. A further purpose of the assignment was to enable students to develop skills in describing research in cognitive psychology to the lay public. The assignment involved reading the chosen book during the course of the semester and keeping a journal of reflections related to the purpose of the assignment. Students also were expected to write a professional style book review (of the type that might appear in the *New York Times* Review of Books). The rubric is abbreviated for inclusion here.

	Unacceptable	Fair	Proficient	Exemplary
<b>Book Journal</b>				
<b>Use of grammar and style to communicate ideas effectively</b>	Grammar and style that interfere with a reader’s ability to understand the ideas presented	Grammar and style adequate for the reader to grasp the main concepts presented	Grammar and style allow the reader to understand easily the concepts presented	Grammar and style enhance the reader’s ability to understand the concepts presented, including nuances of thought; May provide a pleasurable reading experience
<b>Engagement with the author’s ideas</b>	Author’s ideas are simply repeated, indicating that engagement was at or below a surface level	Occasional discussion of the author’s ideas, suggesting ability to engage	Frequent discussion and analysis of the author’s ideas, including expression of well-supported opinions about those ideas, suggesting almost constant engagement	Rich, mature grasp of the author’s ideas, coupled with analysis and synthesis with own ideas and ideas of other writers and scholars, suggesting constant and sophisticated engagement
<b>Connections between the course and the book</b>	Very few connections with course material	Sporadic but meaningful connections with course material	Regular and meaningful connections to course material	Continual connections to course material and sophisticated discussion of those connections
<b>Connections between other experiences and the book</b>	Very few connections with other experiences	Sporadic but meaningful connections with other experiences	Regular and meaningful connections with other experiences	Continual connections to other experiences and sophisticated discussion of those connections

<b>Book Review</b>				
<b>Grammar and form</b>	Grammar and style impede understanding of the “plot” or thesis of the book; not consistent with the form of a professional book review	Grammar and style are adequate for the reader to grasp the “plot” or thesis of the book; the form is consistent with that of a professional book review	Grammar and style allow the reader to understand easily the “plot” or thesis of the book; closely adheres to the style and form of a professional book review	Grammar and style enhance the reader’s ability to understand the “plot” and thesis of the book; indistinguishable from a professional book review
<b>Communication of cognitive psychology concepts to the reader</b>	Ignores the reader’s perspective of the reader and/or communicates cognitive psychology concepts inaccurately or without scientific analysis	Sometimes considers the perspective of the reader and occasionally communicates cognitive psychology concepts well	Consistently addresses the perspective of the reader and communicates cognitive psychology concepts accurately and usefully	Engages the reader and “forces” the reader to be interested in the topic of the book; describes cognitive psychology concepts accurately and usefully

## Test Blueprints

The creation of local examinations—“traditional” examinations at the course level, or comprehensive examinations at the program level—ideally begins by writing a test blueprint before developing the actual test questions. Often called a table of specifications, a test blueprint is a list of the key learning outcomes to be assessed on the test, with the number of points or test questions to be devoted to each goal.

An example of a test blueprint appears in Figure 11. Note that in a test blueprint, an essential learning outcome might be represented by questions worth a total of 20 points, while a lesser learning outcome might be represented by only 5 points.

The test blueprint itself is important evidence of the test’s validity. When matched with test scores, it offers clear evidence of what students have learned because it covers all learning goals. One could say with confidence, for instance, that a student earning an “A” on the test has mastered all or most of the important learning outcomes for a course or a program.

## Other Assessment Tools

Some other assessment tools may be valuable components of many successful assessment programs, but they are more difficult or time-consuming to implement than the tools suggested above, and they also may require significant financial resources to purchase or administer. Careful consideration is warranted to determine whether information yielded from these strategies justifies the time and effort they require.

### Multidimensional or Comprehensive Tests

As most faculty members are already aware, valid and reliable tests can be very difficult to design, especially those meant to assess higher-order thinking skills, attributes, or values. Tests of this type should be administered, analyzed, and revised over several semesters to eliminate poorly written items and to ensure optimal quality. It is best to seek the advice of a colleague who is an expert in “tests and measurements” before embarking on the construction of a comprehensive test of multiple student learning goals. Several books are primers on test construction. At the very least, they will provide the reader with an overview of the best

Figure 9

## Example of a Holistic Scoring Guide (For Critical Thinking)

by Facione and Facione

[Ed. Note: The criteria below are shown from the highest score to the lowest.]

**4** Consistently does all or almost all of the following:

- Accurately interprets evidence, statements, graphics, questions, etc.
- Identifies the salient arguments (reasons and claims) pro and con
- Thoughtfully analyzes and evaluates major alternative points of view
- Draws warranted, judicious, non-fallacious conclusions
- Justifies key results and procedures, explains assumptions
- Fair-mindedly follows where evidence and reasons lead

**3** Does most or many of the following:

- Accurately interprets evidence, statements, graphics, questions, etc.
- Identifies relevant arguments (reasons and claims) pro and con
- Offers analyses and evaluations of obvious alternative points of view
- Draws warranted, non-fallacious conclusions
- Justifies some results or procedures, explains reasons
- Fair-mindedly follows where evidence and reasons lead

**2** Does most or many of the following:

- Misinterprets evidence, statements, graphics, questions, etc.
- Fails to identify strong, relevant counter-arguments
- Ignores or superficially evaluates obvious alternative points of view
- Draws unwarranted or fallacious conclusions
- Justifies few results or procedures, seldom explains reasons
- Regardless of the evidence or reasons, maintains or defends views based on self-interest or preconceptions

**1** Consistently does all or almost all of the following:

- Offers biased interpretations of evidence, statements, graphics, questions, information, or the points of view of others
- Fails to identify or hastily dismisses strong, relevant counter-arguments
- Ignores or superficially evaluates obvious alternative points of view
- Argues using fallacious or irrelevant reasons, and unwarranted claims
- Does not justify results or procedures, nor explain reasons
- Regardless of the evidence or reasons, maintains or defends views based on self-interest or preconceptions
- Exhibits close-mindedness or hostility to reason

© 1994, Peter A. Facione, Noreen C. Facione, and The California Academic Press.

For further information, contact the authors at Insight Assessment ([info@insightassessment.com](mailto:info@insightassessment.com); Phone: 650-692-5628) or visit the website at <http://calpress.com/rubric.html> for a reproducible version and instructions.

Figure 10

## Student Self-reflection Questions for a Course or Program

1. How do you feel about writing/teaching/biology/sociology/etc.?
2. What will you say to your friends about this course/program?
3. What suggestions would you give other students on ways to get the most out this course/program?
4. How do you feel about yourself as a writer/teacher/biologist/sociologist/etc.?
5. What are your strengths as a writer/teacher/biologist/sociologist/etc.?
6. What makes a person a good writer/teacher/biologist/sociologist/etc.?
7. What was the one most useful or meaningful thing you learned in this course/program?
8. What was your biggest achievement in this course/program?
9. In what area did you improve the most? What improvement(s) did you make?
10. What one assignment for this course/program was your best work? What makes it your best work? What did you learn by creating it? What does it say about you as a writer/teacher/biologist/sociologist/etc.?
11. Describe something major that you have learned about yourself in this course/program.
12. List three ways you think you have grown or developed as a result of this course/program.
13. In what ways have you improved as a writer/teacher/biologist/sociologist/etc.?
14. What have you learned in this course/program that will help you continue to grow as a writer/teacher/biologist/sociologist/etc.?
15. What was your favorite aspect of this course/program? Why?
16. What goals did you set for yourself in this course/program? How well did you accomplish them?
17. If you were to start this course/program over, what would you do differently next time?
18. What strategies did you use to learn the material in this course/program? Which were most effective? Why?
19. What risks did you take in this course/program?
20. If you could change any one of the assignments you did for this course/program, which one would it be?  
  
What would you change about it?
21. What problems did you encounter in this course/program? How did you solve them?
22. What one question about this course/program is uppermost on your mind?
23. What would you like to learn further about this subject/discipline?
24. In what area would you like to continue to strengthen your knowledge or skills?
25. Write one goal for next semester/year and describe how you plan to reach it.



Figure 11

## Example of a Test Blueprint

### Educational Research Methods: Final Exam Outline

The final exam will consist of 25 multiple-choice items, each worth 2 to 4 points, and five short essay questions, each worth 3 to 5 points. The items will cover most of the concepts listed below.

#### ***Validity and Reliability (Up to 16 points)***

- Demonstrate an understanding of reliability and validity.
- Correctly identify the type of reliability and validity evidence being provided by given information on an instrument.
- Recognize examples of measurement error in a given situation.
- Assess the meaning and implications of measurement error.
- Apply general principles for ensuring validity.

#### ***Inferential Statistics (Up to 16 points)***

- Demonstrate an understanding of the concept of a null hypothesis.
- Select the most appropriate inferential statistics (t, F, or  $\chi^2$ ) for a given research situation
- Know the most common “cut-off” point that statisticians use in deciding whether two means differ statistically significantly from one another.
- Correctly interpret the results of t, F, and  $\chi^2$  tests as presented in research articles.
- Interpret the effect of standard deviation and sample size on the results of a statistical test.

#### ***Experimental Research (Up to 12 points)***

- Interpret correctly the symbolic representations of experimental designs.
- Describe the benefits and limitations of each experimental and quasi-experimental design covered in class.
- Identify the appropriate research design for a given research situation.

#### ***Correlational Research (Up to 12 points)***

- Demonstrate an understanding of regression and the use of regression equations.

- Understand what r,  $R^2$ , and partial correlations are and what they tell us.
- Understand what multiple regression analysis is used for and what it tells us.

#### ***Qualitative Research: Observation, Interviews, and Ethnographic Research (Up to 16 points)***

- Describe and discuss qualitative research and its key characteristics.
- Identify the pros and cons of qualitative research.
- Describe the concept of a focus groups.
- Identify the pros and cons of focus group research.
- Understand the key principles in conducting focus groups.
- Define ethnographic research is and identify or describe examples of it.

#### ***Historical Research (Up to 10 points)***

- Describe the need for historical research.
- Identify kinds of historical research sources.
- Recognize examples of primary and secondary resources.
- Understand how to evaluate historical research.

#### ***Content Analysis (12 points)***

- Demonstrate an understanding of content analysis.
- Understand the pros and cons of content analysis.
- Recognize examples of different kinds of content analysis.
- Explain how to analyze content analysis data.

#### ***Multiple Units (Up to 6 points)***

- Identify the most appropriate research method for a given situation.

questions to ask when seeking expert advice (Anastasi and Urbina, 1996; Haladyna, 1999).

Adding a published test to an assessment program will require time to identify, evaluate, and experiment with potential tests. Unfortunately, many published tests aimed at the higher education market offer limited evidence of quality (i.e., validity and reliability) and have been normed with relatively small groups of students.

It is most important to compare the test blueprint against the key learning outcomes of the course or program in question to see how well they match. A biology test that focuses on ecological concepts, for example, probably would not be appropriate as a key assessment instrument for a biology program that aims to prepare students for careers in health professions.

Figure 12 contains a list of published tests designed to test critical thinking and general education goals. It is presented here as an example of the various test characteristics that should be considered when choosing an appropriate published assessment instrument.

### ***Ad Hoc* Surveys and Pre-graduation Surveys**

Many people view surveys as a quick way to collect assessment information. Unfortunately, surveys that are designed and administered quickly often have low response rates and poorly-phrased questions that yield information of questionable value.

Indirect assessments of student perceptions and satisfaction that are administered at the institutional level and are not embedded in course and program requirements—such as voluntary graduating senior surveys—take extra time and effort for both students and faculty members, and they present sampling problems. It also can be difficult to motivate students to participate in such extraneous assessment efforts, or to give their best possible effort and thought, thus reducing the validity of the assessment itself. It is often simpler, more efficient, and more effective to use assessment strategies that

are intrinsic parts of course and program requirements. Graduating senior surveys, for instance, could be administered as part of a capstone course offered in every major.

If an institution determines that a survey is a key element of an assessment strategy, it should help to ensure useful results by conducting a pilot test of the survey. A draft should be administered to a small group, the responses analyzed, and unclear questions identified. Strategies to maximize the response rate should be included in the plans to administer the actual survey.<sup>5</sup>

### **Focus Groups**

A focus group interview often is viewed as another quick way to collect assessment information, but the relatively small number of participants and the free-flowing format can reduce the credibility and value of the results. Focus groups are usually most appropriate as tools to help illuminate other assessment results, rather than as stand-alone assessment strategies.

Successful focus groups require time for planning, testing, and analysis to ensure a balanced discussion among a sufficient number of participants and to assure that the results have credibility and value. One should learn how to plan and conduct focus groups, hire a consultant, or enlist the aid of an on-campus expert before using focus groups as an assessment strategy.

Several sources introduce the science of conducting focus groups and their use as a source of information. For example, see Morgan (1997); and Krueger and Casey (2000).

### **Portfolios**

Portfolios are structured, focused, and purposeful collections of student work. They are increasingly popular assessment strategies, because they provide an exceptionally comprehensive, holistic picture of student learning.

Figure 13 offers some questions that may help in a decision on whether or not to use portfolios. If the decision is made to use portfolios, it is best to

---

<sup>5</sup> For a discussion of effective survey use, see Suskie (1996).

Figure 12

## Commonly-administered Measures of Critical Thinking

Measure	Critical Thinking Definition	Subscales	Design	Appropriate Participants
<b>Watson-Glaser Critical Thinking Appraisal</b>	Comprises attitudes, knowledge, skills	Inference, Recognition of assumptions, Deduction, Interpretation, Evaluation of arguments	Parallel forms A & B; 80 multiple-choice items, based on readings; 40 mins. to complete	9th grade and higher
<b>California Critical Thinking Skills Test</b>	Purposeful, self-regulatory judgment	Analysis, Evaluation, Inference, Inductive, Deductive	Parallel forms A and B; 34 items; 40 mins.to complete	College age
<b>California Critical Thinking Dispositions Inventory</b>	Attitudinal inclination to apply critical thinking skills	Truth seeking, Open mindedness, Analyticity, Systematicity, Critical thinking self-confidence, Inquisitiveness, Cognitive maturity	Likert-type scale; 75 items; Response ranges from Agree to Strongly Disagree; 40 mins.to complete	College age
<b>Ennis-Weir Critical Thinking Essay Test</b>	Reasonably deciding what to do or what to believe	Getting the point; Seeing reasons and assumptions; Stating one's point; Offering good reasons; Seeing other possibilities; Equivocation; Irrelevance; Circularity; Reversal of conditional relationships; Straw person fallacy; Overgeneralizations; Excessive skepticism; Credibility; Using emotive language to persuade	Essay format; Responses written to questions about scenarios; 40 minutes to complete	Grade 7 to College
<b>Cornell Critical Thinking Test</b>	Reasonably deciding what to do or believe	Level X: Induction; Deduction. Credibility, Assumptions, Value judgment; Meaning Level Z: All level X subscaled plus semantics, prediction, definition	Level X: 71 multiple-choice items based on scenarios; Level Z: 52 multiple-choice items based on scenarios; 50 mins. to complete	Level X: 4th grade-college sophomore  Level Z: gifted high school and college-aged adults

References: Adams, M., Whitlow, J., Stover, L., and Johnson, K. (1996). Critical thinking as an educational outcome: An evaluation of current tools of measurement. *Nurse Educator*, 21 (3), 23-31.

Facione, N.C. (1997). *Critical thinking assessment in nursing education programs: An aggregate data analysis*. Millbrae: The California Academic Press.

Ennis, R. H. & Millman, J. (1985). *Cornell critical thinking test, Level 2*. Pacific Grove, California: Midwest Publications.

Ennis, R. H. & Millman, J. (1985). *Cornell critical thinking test, Level X*. Pacific Grove, California: Midwest Publications.

Rane-Szostak, & D. Robertson, J. F. (1996). Issues in measuring critical thinking: Meeting the challenge. *Journal of Nursing Education*, 35(1), 5-11.

Note: This table contains information on the most commonly used measures of critical thinking only. It is not meant to be exclusive. There are many more measures available, including several domain-specific measures. Table prepared in 2001 by D. A. Redding, Ph.D., Instructional Assistant Professor, Mennonite College of Nursing, Illinois State University. Reproduced with permission..

start on a small scale. Portfolios may be especially appropriate for programs that enroll only a handful of students. Such programs would be ideal for piloting portfolio projects for later use with larger programs.

Portfolios can present significant logistical problems related to sampling, storage, development of evaluation criteria, and the allotment of sufficient faculty and staff time for review. These issues can be resolved, but the solutions may take time to identify and implement. For example, a number of institutions use electronic portfolios to solve the storage problem. Huba and Freed (2000) provide an excellent discussion of the development and assessment of portfolios.

### **Retention/Graduation Rates**

Retention and graduation rates that do not meet the institution's goals may be signs of problems with student learning. However, they do not necessarily reveal what students actually have learned. They can be useful to the extent that they correlate with and illuminate direct learning assessments, or that they assess directly such institutional outcomes as cost effectiveness, diversity, student achievement, and other evaluations of institutional effectiveness.

Figure 13

## **Considerations when Deciding to Use Portfolios**

1. What are the goals of the portfolio?
  - What do you want your students to learn by the act of creating a portfolio?
  - What processes or outcomes are to be evaluated by the portfolio?
2. How will students choose what to include in the portfolio?
3. How and when will work be included in the portfolio?
4. How will student and faculty reflection occur in the portfolio process?
5. How will the portfolios be reviewed and evaluated? What would a successful portfolio in your program look like? What are your criteria for deciding if a portfolio is a "success"?
6. Will the portfolios be graded? If so, how?
7. How and where will portfolios be stored?
8. Will the portfolios be passed one faculty member to another? Will students retain ownership of portfolios?
9. What are the benefits of moving toward portfolio assessment? What are the areas of concern?
10. Is the collection of student work a feasible practice in your program?

© 2002, Copyright by Linda Suskie. Reproduced with permission.

# 4

## Understanding Middle States Expectations for Assessing Student Learning and Institutional Effectiveness

In 2002, the Middle States Commission on Higher Education introduced updated accreditation standards that simplified requirements for resources and processes and concentrated instead on *assessment*: evidence that the institution is achieving its goals. Every accreditation standard now includes an assessment component; the assessment of student learning is addressed in Standard 14 (Assessment of Student Learning); and the assessment of *all* key institutional goals, including those assessed in the other thirteen standards, is addressed holistically in Standard 7 (Institutional Assessment).

Because Standards 7 and 14 are a significant change from prior standards, and because the Commission gives institutions great latitude in choosing approaches to comply with them, these two standards have engendered many questions. This statement is intended to address these questions and to clarify the Commission's expectations regarding these standards and their relationship to other standards such as Standard 2 (Planning, Resource Allocation, and Institutional Renewal).

### *What is the Assessment of Institutional Effectiveness (Standard 7)?*

Assessment may be characterized as the third element of a four-step planning-assessment cycle:

1. Defining clearly articulated institutional and unit-level *goals*;
2. Implementing *strategies to achieve those goals*;
3. *Assessing achievement* of those goals; and
4. *Using the results* of those assessments to improve programs and services and inform planning and resource allocation decisions.

The effectiveness of an institution rests upon the contribution that each of the institution's programs and services makes toward achieving the goals of the institution as a whole. ***Standard 7 (Institutional Assessment) thus builds upon all other accreditation standards, each of which includes periodic assessment of effectiveness as one of its fundamental elements.*** This standard ties together those assessments into an integrated whole to answer the question, "As an institutional community, how well are we collectively doing what we say we are doing?" and, in particular, "How do we support student learning, a fundamental aspect of institutional effectiveness?" (Standard 14). Self-studies can thus document compliance with Standard 7 by summarizing the assessments within each accreditation standard into

conclusions about the institution's overall achievement of its key goals.

### ***What is the Assessment of Student Learning (Standard 14)?***

Assessment of student learning may be characterized as the third element of a four-step teaching-learning-assessment cycle that parallels the planning-assessment cycle described above:

1. Developing clearly articulated *learning outcomes*: the knowledge, skills, and competencies that students are expected to exhibit upon successful completion of a course, academic program, co-curricular program, general education requirement, or other specific set of experiences;
2. Offering courses, programs, and experiences that provide purposeful *opportunities for students to achieve those learning outcomes*;
3. *Assessing student achievement* of those learning outcomes; and
4. *Using the results* of those assessments to improve teaching and learning and inform planning and resource allocation decisions.

***Because student learning is a fundamental component of the mission of most institutions of higher education, the assessment of student learning is an essential component of the assessment of institutional effectiveness (Standard 7) and is the focus of Standard 14 (Assessment of Student Learning).***

### ***Why Does the Commission Expect Student Learning and Institutional Effectiveness to be Assessed?***

The fundamental question asked in the accreditation process is, "Is the institution fulfilling its mission and achieving its goals?" This is precisely the question that assessment is designed to answer, making assessment essential to the accreditation process. Assessment processes help to ensure that:

- ❑ Institutional and program-level goals are clear to the public, students, faculty, and staff.
- ❑ Institutional programs and resources are organized and coordinated to achieve institutional and program-level goals.
- ❑ The institution is indeed achieving its mission and goals.
- ❑ The institution is using assessment results to improve student learning and otherwise advance the institution.

### ***What Are the Characteristics of Assessment Processes that Meet Middle States Expectations?***

Effective assessment processes are useful, cost-effective, reasonably accurate and truthful, carefully planned, and organized, systematic, and sustained.

1. **Useful** assessment processes help faculty and staff make appropriate decisions about improving programs and services, developing goals and plans, and making resource allocations. Because institutions, their students, and their environments are continually evolving, effective assessments cannot be static; they must be reviewed periodically and adapted in order to remain useful.
2. **Cost-effective** assessment processes yield dividends that justify the institution's investment in them, particularly in terms of faculty and staff time. To this end, institutions may begin by considering assessment measures, indicators, "flags," and "scorecards" already in place, such as retention, graduation, transfer, and placement rates, financial ratios, and surveys. New or refined measures may then be added for those goals for which evidence of achievement is not already available, concentrating on the institution's most important goals. Effective assessments are simple rather than elaborate, and they may focus on just a few key goals in each program, unit, and curriculum.

3. **Reasonably accurate and truthful** assessment processes yield results that can be used with confidence to make appropriate decisions. Because there is no one perfectly accurate assessment tool or strategy, institutions should use multiple kinds of measures to assess goal achievement. Assessments may be quantitative or qualitative and developed locally or by an external organization. All assessment tools and strategies should clearly relate to the goals they are assessing and should be developed with care; they should not be not merely anecdotal information nor collections of information that happen to be on hand. Strategies to assess student learning should include *direct*—clear, visible, and convincing—evidence, rather than solely indirect evidence of student learning such as surveys and focus groups.

4. **Planned** assessment processes that are purposefully linked to institutional goals promote attention to those goals and plans and ensure that disappointing outcomes are appropriately addressed. Institutions often have a variety of plans, such as a strategic plan, academic plan, financial plan, enrollment plan, capital facilities master plan, and technology plan. Just as such plans should be interrelated to ensure that they work synergistically to advance the institution, assessments should also be interrelated. At many institutions, effective institutional planning begins with academic planning, which in turn drives the other plans. If the academic plan calls for a new academic program, for example, the technology plan should ensure faculty and students in the new program will be able to use appropriate instructional technologies. Assessments of the technology plan should evaluate not just whether instructional technologies have been put in place but also how effectively those technologies have helped students to achieve the program’s key learning outcomes.

5. **Organized, systematized, and sustained** assessment processes are ongoing, not once-and-done. There should be clear interrelationships among institutional goals, program- and unit-level goals, and course-level goals.

## ***What Should Institutions Document Regarding Assessment?***

When submitting information on their assessment efforts to the Commission, institutions are expected to document:

- clear statements of key goals, including expected student learning outcomes;
- an organized and sustained assessment process (referred to in some Commission documents as an “assessment plan”) including:
  - institutional guidelines, resources, coordination, and support for assessment;
  - assessment activities and initiatives that are presently underway;
  - plans to develop and implement future assessment activities and initiatives;
- assessment results demonstrating that the institution and its students are achieving key institutional and program goals; and
- uses of assessment results to improve student learning and advance the institution.

## ***How Should This Information Be Organized and Formatted for Review by the Commission and its Representatives?***

Assessment documentation that is organized into a coherent presentation of what the institution is doing regarding assessment provides a roadmap that facilitates the work of evaluation teams, reviewers, and the Commission. Assessment documentation is typically a living, fluid, organized collection of documents and/or online resources, often with references and/or links to further documents and online resources, that are routinely updated as the institution’s assessment processes evolve. There is not, however, any prescribed format or organization for these materials; institutions have maximum flexibility in designing and assembling assessment documentation that fits best with the institution’s mission, organization,

and needs. *A single, formal, polished document is not required* and, for many institutions, may not be the most suitable format, because it may discourage the continual modifications that are made in effective assessment processes. The existence of an effective process, clearly described to the community and the Commission, is more important than a formal plan.

Institutions may choose to include an appropriate combination of the following in their assessment documentation

- ❑ **An overview in a self-study, periodic review report, or follow-up report** gives the Commission and its representatives a useful introductory synopsis of the institution's assessment processes.
- ❑ **A chart or "roadmap" outlining assessment documentation**, provided within a self-study or periodic review report or as an appendix, can be especially useful for large or complex institutions with a broad array of goals and assessment processes.
- ❑ **A written or online assessment plan** that documents an organized, sustained assessment process (including institutional guidelines, resources, coordination, and support for assessment, assessment activities and initiatives that are presently underway, and plans to develop and implement future assessment activities and initiatives) can be an excellent way to initiate, structure, and demonstrate compliance with Standards 7 and 14, although it is not required. Assessment plans can guide and support the institutional community in its efforts to assess its mission and goals by:
  - helping to ensure that assessment is efficient, effective, and purposeful, rather than just a collection of available information,
  - providing information needed to carry out assessment practices, and
  - helping to ensure that assessment is supported with appropriate resources and that results are used appropriately.

- ❑ **Assessment documentation incorporated within the institutional (strategic) plan** or in separate documentation clearly linked to the institutional plan.
- ❑ **Separate assessment documentation for each institutional division** that is linked together may be a feasible approach, especially for large, complex institutions.
- ❑ **More thorough information in an on-site resource room and/or online** enables evaluation team members to review a cross-section of program- and unit-level assessment processes.

### *How Are the Documentation of Institutional Assessment and Student Learning Assessment Related?*

As noted earlier, because student learning is a fundamental component of the mission of most institutions of higher education, the assessment of student learning is an essential component of the assessment of institutional effectiveness. An institution may therefore create institutional effectiveness documentation that includes a component on assessing student learning, or it may create a bridge between two separate sets of documentation, one for the assessment of student learning and one for other aspects of institutional effectiveness.

### *What Might the Commission and Its Representatives Look For in Assessment Documentation?*

Evaluation team members, reviewers, and Commissioners might look for information on the following questions in an institution's assessment documentation:

1. **Do institutional leaders support and value a culture of assessment?** Is there adequate, ongoing guidance, resources, coordination, and support for assessment? (This may include administrative support, technical support, financial support, professional development, policies and procedures, and governance structures that ensure appropriate



collaboration and ownership.) Are assessment efforts recognized and valued? Are efforts to improve teaching recognized and valued?

2. **Are goals, including learning outcomes, clearly articulated at every level:** institutional, unit-level, program-level, and course-level? Do they have appropriate interrelationships? Do the undergraduate curriculum and requirements address institutional learning outcomes and the competencies listed in Middle States' Standard 12 (General Education)? Are all learning outcomes of sufficient rigor for a higher education institution? Are learning outcomes for, say, master's programs more advanced than those for undergraduate programs?

3. **Have appropriate assessment processes been implemented** for an appropriate proportion of goals? (Expectations for an "appropriate proportion" are increasing.) Do they meet Middle States expectations, as characterized above?

4. Where assessment processes have not yet been implemented, **have appropriate assessment processes been planned?** Are the plans feasible? Are they simple, practical, and sufficiently detailed to engender confidence that they will be implemented as planned? Do they have clear ownership? Are timelines appropriate, or are they either overly ambitious or stretched out too far?

5. **Do assessment results provide convincing evidence** that the institution is achieving its mission and goals, including key learning outcomes?

6. **Have assessment results been shared** in useful forms and discussed widely with appropriate constituents?

7. **Have results led to appropriate decisions** and improvements about curricula and pedagogy, programs and services, resource allocation, and institutional goals and plans?

8. **Have assessment processes been reviewed** regularly? Have the reviews led to appropriate decisions and improvements in assessment processes and support for them?

9. **Where does the institution appear to be going with assessment?** Does it have sufficient engagement and momentum to sustain its assessment processes? Or does it appear that momentum may slow? Are there any significant gaps in assessment processes, such as key areas where no assessment plans have been developed?

# 5

## Using Assessment Results To Improve Teaching and Learning

**A** commitment to the assessment of student learning requires a parallel commitment to ensuring its use. Perhaps the most difficult part of assessing student learning is the process of effecting change in teaching and learning as a result of information gained through assessment practices. It is pointless simply to “do assessment”; the results of assessment activities should come full circle to have a direct impact on teaching and learning and on the institution’s strategic plan to fulfill its mission.

Continuous improvement can occur in an upward spiral if an institution’s structure is flexible, and if members of the campus community are committed to the assessment plan and are willing to integrate the results of assessing student learning into their collective vision of what the institution is doing well and what it could do better.

The first section of this chapter discusses the ways in which institutions can encourage the use of assessment results, the second section presents examples of specific types of change that might be made as the result of information gained in assessment, and the third section discusses the interconnectedness of assessment, teaching, and learning.

### *Institutional Support Strategies Designed to Encourage the Use of Assessment Results*

An assessment plan will serve its purpose only if it provides for the use of assessment results. Regardless of the level at which assessment is conducted, an articulated plan for translating assessment results into changes in practice is essential. For such a plan to be effective, it requires an institutional commitment to the use of assessment results, the sharing of results, a broad campus discussion of and decision-making on those results, individuals who are empowered to make changes, the availability of resources for change, and flexible procedures for implementing changes.

#### **An Institutional Commitment**

The institution should demonstrate a commitment to developing a system for analyzing results, identifying areas of strength and weakness, creating a strategy for improving the learning experience, and implementing that strategy. Such a commitment will increase student learning as well as increase faculty and staff commitment to assessment. However, if the results of assessment are not used to improve student learning, assessment becomes at best a descriptive set of data about students and, at worst, a useless exercise.

Consider a business department that collects data regarding student or alumni performance on The American Institute of Certified Public Accounting Uniform CPA Examination and discovers that the majority of its graduates are failing the exam. This knowledge provides the opportunity to review the relevant parts of the curriculum, implement strategies for change, and gauge any improvement in student learning as a result of the changes made. In contrast, a tacit decision not to make curricular changes after discovery of this information could result in the demoralization of students and faculty, diminished stature for the program, and reduced selectivity in admissions.

Changes in programmatic curricula as a result of assessment data do not happen automatically, as many faculty and staff members can attest. However, if the department plan outlines specific procedures for examining assessment results and implementing curricular revision, those changes are more likely to occur.

### **Sharing Assessment Results**

Assessment data collected at the institutional and program levels should be made available to the relevant members of the campus community. Data at the course level should be shared when it is appropriate to do so, such as when faculty members are collaborating to develop or revise a course, or are team-teaching a course.<sup>6</sup> When assessment data are collected but not shared with those who would be responsible for implementing change, the data are useless for practical purposes. Similarly, a perceived lack of faculty interest in assessment could be caused by the belief that assessment initiatives yield little or no meaningful information.

The first problem—when data are collected but not shared with those responsible for implementing change—can occur when one area or program collects data that are relevant to another area but fails to make the data available. For instance, social science faculty may assess their students' research performance via a required common paper, presentation, or capstone course. Assessments might reveal that students are not achieving desired

levels of information literacy. Students may fail to use analytical thinking when critiquing primary source articles, may cite materials improperly, or may conduct inadequate literature searches. This information can help in revising social sciences courses, but it also would be of great value to library staff members who design and deliver significant components of the information literacy requirement.

The second problem—when faculty members show little interest in assessment because they perceive it as meaningless—can result when data are collected at the institutional level to satisfy an outside agency, such as a state board of education or an accreditor, but are never shared with the faculty. In such cases, the failure to share data may be the result of hectic and unplanned-for data collection rather than an intentional withholding of important information from campus stakeholders. If there is no planned provision for collecting assessment data—for instance, if they are collected *ad hoc* to satisfy an external agency—there is unlikely to be a provision to share them regularly with the campus community.

There are cases in which an institution decides not to share data because it fears that assessment results indicating that students are not achieving desired levels of learning or that students are not satisfied will be shared with the general public and will impair the institution's ability to attract students. This is counter-productive for several reasons: silence by the institution about student performance is itself a red flag to the public, and poor performance by the institution's graduates will nevertheless be noticed by employers and the public. Most importantly, failure to share information with internal stakeholders precludes the opportunity to improve and to produce the type of student learning that will attract students to the institution. Even if an institution chooses justifiably not to publicize certain results externally, it should ensure that useful data are shared and used internally.

---

<sup>6</sup> It is advisable to separate the assessment of student learning from the assessment of an individual faculty member's teaching, wherever possible, in order to encourage faculty members to engage in assessment activities, not to shun them.

For example, many institutions exercise their prerogative not to disclose to the public their results from the National Survey of Student Engagement (NSSE). The NSSE is a student self-report survey that measures what it describes as “student participation in programs and activities that institutions provide for their learning and personal development. The results provide an estimate of how undergraduates spend their time and what they gain from attending college.”<sup>7</sup> In the NSSE, students report the number of hours they spent on schoolwork outside of class, the number of written papers completed, and the length of their papers, in addition to many other aspects of how they spend their time. Results such as these about student engagement can be invaluable to faculty members, librarians, and student affairs staff as they revise curricula and programs. Therefore, administrators and faculty members should give priority to the best interests of students by devising a system to share even sensitive assessment results internally, regardless of the test or measure from which they resulted.

### **Campus Discussion and Shared Decision-making**

Assessment results are less likely to produce meaningful improvement in learning if only a small number of people or offices make all of the decisions about modifications to the learning experience.

Students should be included in discussions about assessment whenever possible, and they should be encouraged to engage in conversations with their peers about the institution’s curricula and programs. Many campuses have specific courses or other learning activities that become the nexus of student complaints. For example, some general education courses frequently become the focus of complaints about a lack of “real world meaning” and connection to the major. Discussions about assessment results and curricular modification are an ideal venue to channel students’ comments and criticisms constructively.

### **Empowering Individuals to Effect Change**

Clear and public charges should be made to those who will be responsible for leading programmatic and curricular change that occurs as a result of assessment. At the course level, the individual instructor or group of instructors who teach a specific course would, of course, be responsible for its revision. At the program level, someone such as the department or program chair may be given the responsibility to ensure that change occurs. This person is often the same person who implemented the assessments. At the institutional level, however, several people from across the institution will be responsible for assessing and for changing the curriculum. For instance, the Office of Institutional Research might collect the data, and other offices or departments may be charged with effecting change.

It is important to articulate exactly who is responsible for change so that the data do not stagnate “on the shelf.” For example, even if the office of career services is charged with conducting an annual survey on student acceptance rates at graduate and professional schools, it should be made clear which faculty members are responsible for implementing programs to improve education and increase graduate school acceptance rates.

### **Resources for Change and Celebration of Achievements**

After assessment data are collected and curriculum and program revisions have been planned, resources must be available to implement the changes. Unfortunately, funds often are not available for every suggested change. Faculty members and administrators should review the institution’s mission and strategic plan to determine funding priorities for new initiatives and to weigh the costs and benefits of proposed changes. A clear process for determining budgetary priorities should ensure commitment to the best interests of all students, rather than giving priority to the interests of a small group of faculty or students.

---

<sup>7</sup> See [www.iub.edu/~nsse/html/facts.shtml](http://www.iub.edu/~nsse/html/facts.shtml) .

Assessment successes need to be positively reinforced in a way that makes the campus community aware of the value of assessment. Yearly celebrations can focus on effective assessment strategies, positive change as a result of assessment, or new assessment ideas. More importantly, traditional reward systems related to faculty evaluation, promotion, and tenure should take into account the valuable work of assessment.

### **Flexibility**

This handbook has stressed the importance of well-communicated and clear procedures and plans for developing and implementing assessment programs. *Procedures, however, need not be arduous or cumbersome to result in positive change.* Inflexible and bureaucratic procedures discourage faculty from embracing assessment and adapting courses and programs in response to assessment results. For instance, engineering technology faculty members might decide that students need mini-practicum experiences early in their undergraduate careers because general student performance on senior projects is inadequate. Faculty are much more likely to modify the program to include early practicum experiences if their proposal is not stalled in committees. Institutions should strive to develop facilitative procedures that include relevant stakeholders and do their best to avoid bureaucratic structures that discourage change.

### ***Translating Assessment Results into Better Learning***

The most important reason for assessment is to ensure that students are learning. Even when the requirements of those to whom the institution is externally accountable—students, parents, legislators, and accreditors—provide the impetus for assessment, the fundamental expectation is that institutions of higher learning demonstrate that their students are learning.

Unfortunately, there are many obstacles to change. Faculty often object to performing yet another task related to assessment, citing additional demands on their time. They also might believe that the results of some assessment activities are invalid, or that the results demonstrate merely what the administration

wants them to demonstrate. Alternatively, institutions and committees may exhaust themselves planning for assessment and become “burned out” before results actually affect learning. Even when faculty members are committed to using assessment results to improve learning, the institution may not commit the necessary resources. It is common for accreditation teams to find beautiful assessment plans that have yet to be enacted, or “completed” assessment plans for which the resultant data sit on the shelf because the institution has not committed sufficient human or economic resources to support change.

Using assessment results need not be an onerous task, particularly for faculty who regularly adapt and modify their courses for the sake of their students. Using assessment results means changing courses or programs on the basis of real data rather than intuition. Even seasoned professors might be surprised by assessment data. Perhaps students are not reading a text because it is too elementary for them or too factual, instead of providing the type of analysis that might inspire their interest. Students who are performing extremely well on examinations nevertheless may not have been sufficiently challenged by the course. Perhaps students are performing poorly on one type of examination (e.g., an essay) because the mode of teaching was more conducive to performing well on another type of examination (e.g., multiple choice). The causes of ineffective learning experiences cannot always be explained by intuitive hunches.

Figures 15, 16, and 17 describe hypothetical and diverse ways of using assessment results at the institutional, course, and program levels. Although they are presented in a condensed and oversimplified form, they are not intended to imply that changing curricula, programs, or courses is simple or requires little thought. Rather, they are presented in the spirit of encouraging discourse among all members of the campus community, especially among faculty members and students. Each of the examples provides: (1) the initial learning goal at the institutional, course, and program level; (2) the measures or methods of assessing the goal; (3) the outcome evidenced by the measures; (4) the possible reason for the outcome; and (5) the action or actions taken as a result of the assessment.

Several noteworthy points are presented in these examples. First, the reader will notice that not all of the outcomes are “bad.” Some of the examples indicate outcomes such as “students’ projects demonstrate consistently high levels of quality.” Examples of positive outcomes are used here to stress the point that when outcomes are “good”—that is, when students are meeting a learning goal—faculty members should consider whether the grade or the assessment reflect true competence, or whether the assessment or the goal is inappropriately easy and should be more ambitious. However, most of the examples here involve “bad” outcomes, because they provide an opportunity to think about the causes of poor performance and productive changes, not because outcomes in the “real world” are usually “bad.”

The examples in Figures 14, 15, and 16 also illustrate the use of multiple assessments, direct and indirect measures, self-report and “objective” measures, and qualitative and quantitative data. Both direct and indirect measures are presented in the second column for many of the examples to illustrate the importance and interrelation of each. For instance, one of the examples presented in Figure 14 involves one direct measure (standardized writing tests) and three indirect measures (course registration statistics, transcript analysis, and course content analysis).<sup>8</sup> Other variations in form of measurement are represented in these examples as well. For instance, an institution seeking to produce students who compete well in the job market might use qualitative data (e.g., an alumni survey with open-ended questions) and quantitative data (an alumni survey with numerically-scaled questions, together with benchmarking statistics) to assess its goals. Faculty members may use quantitative and objective data (exam scores) and qualitative and self-report data (teaching evaluation comments) to assess the goal that students will master discipline-specific material. Most of these examples use multiple measures as well.

Finally, the “possible reason” or hypothesis for the outcome is presented as if it comes after the “measure” in time. This is not always the case. Often those who are focusing on the enhancement of student learning have developed several hypotheses before the assessments are administered; the assessment measures are used specifically to test the hypotheses. In these examples, the assumption is made that the measures listed are ongoing assessments that are part of a plan. The possible reasons for the outcome or hypotheses about what needs to be changed are drawn either from data gathered from the measures presented in the second column or from clearly related program or course characteristics—not at random or from intuitive suppositions.

For instance, one example in Figure 14 depicts two possible reasons for student dissatisfaction with a first-year community service requirement: the first is students’ failure to understand the relevance to their chosen career or major, and the second is the time and transportation hardships students perceive. Both of these possible reasons for the outcome are drawn from data gathered in the first-year experience student satisfaction survey and focus groups, presented as measures in the second column. Another example, in Figure 15, describes students as having low fluency in modern language after having taken several courses. The hypothesis that lack of sustained and regular practice (i.e., infrequent class meetings) can be drawn from external research about optimum pedagogical practices in foreign language learning; it need not be drawn from a locally administered survey or test. Finally, the reader will notice that occasionally the “action taken” is to collect more data.

Appendix 7 offers a companion brainstorming exercise for which no likely causes or suggested actions have been specified. This exercise can be used with groups who are very new to assessment concepts. It allows for participants to draw on their own impressions and experiences in order to suggest how the faculty members and staff at the hypothetical institutions might proceed.

---

<sup>8</sup> For an in-depth discussion of direct and indirect measures, their uses, and their advantages and disadvantages, see Chapter 3.

Figure 14

## Using Assessment Results at the Institutional Level: Maintaining Mission and Achieving Vision

Learning Goal: Students will...	Direct & Indirect Measures	Outcome	Possible Reason or Hypothesis	Action Taken
<b>Appreciate the importance of civic responsibility</b>	<ul style="list-style-type: none"> <li>* First-year experience student satisfaction survey</li> <li>* Focus groups</li> </ul>	Students express strong dissatisfaction with a first-year community service requirement.	<ul style="list-style-type: none"> <li>* Students do not see relevance to chosen career or major.</li> <li>* Students have time or transportation constraints.</li> </ul>	<ul style="list-style-type: none"> <li>* Introduce student/alumni-run seminars about how community service was relevant for them.</li> <li>* Include references and examples related to community service in general education courses.</li> <li>* Provide transportation, offer credit or work-study funds.</li> </ul>
<b>Exhibit competitive career potential</b>	<ul style="list-style-type: none"> <li>* Alumni survey</li> <li>* Data from benchmark institutions</li> </ul>	Graduates' salaries five years post graduation are lower than those of students with comparable positions who attended other institutions.	<ul style="list-style-type: none"> <li>* Students are not trained in salary negotiation.</li> <li>* Students are under-prepared relative to other institutions.</li> </ul>	<ul style="list-style-type: none"> <li>* Offer opportunities for mock interview and mock raise requests.</li> <li>* Survey employers to determine cause.</li> <li>* Change curricula as a result of employer survey.</li> </ul>
<b>Complete the first-year successfully</b>	<ul style="list-style-type: none"> <li>* First-year retention rates</li> <li>* General education course grades</li> <li>* First-year-experience student satisfaction survey</li> <li>* Exit interviews for non-returning students</li> </ul>	A first-year retention problem is traced to poor performance in general education courses.	Poor writing or analytic skills of students entering the program impede performance.	Change course sequence to offer writing, critical thinking course first.
<b>Complete the Honors Program successfully</b>	<ul style="list-style-type: none"> <li>* Honors program retention figures</li> <li>* Analysis of course availability and registration statistics</li> <li>* First-year-experience survey</li> <li>* Focus groups</li> </ul>	An honors program has low internal retention.	<ul style="list-style-type: none"> <li>* Not enough honors courses are available.</li> <li>* Courses are not judged to be significantly different from regular courses, students do not feel challenged.</li> </ul>	Allow first- and second-year honors students to take upper-division honors courses for credit.
<b>Exhibit independent learning</b>	<ul style="list-style-type: none"> <li>* National Survey of Student Engagement (NSSE) results</li> <li>* Student, faculty, and student affairs staff focus groups</li> </ul>	National Survey of Student Engagement indicates that students spend very little time on academic work outside the classroom.	<ul style="list-style-type: none"> <li>* Students work many hours per week for pay.</li> <li>* Students are not being challenged academically.</li> </ul>	<ul style="list-style-type: none"> <li>* Allow paid internships in area directly relevant to curriculum.</li> <li>* Revamp course syllabi to require more meaningful work outside of class.</li> </ul>

Learning Goal: Students will...	Direct & Indirect Measures	Outcome	Possible Reason or Hypothesis	Action Taken
<b>Demonstrate high-level writing skills</b>	<ul style="list-style-type: none"> <li>* Standardized writing tests in first and third years</li> <li>* Course registration statistics</li> <li>* Transcript analysis</li> <li>* Course content analysis</li> </ul>	Little difference between performance of first-year students and juniors on a standardized writing test.	<ul style="list-style-type: none"> <li>* Students avoid elective courses with a large writing component.</li> <li>* Students have infrequent opportunities to receive feedback on their writing.</li> </ul>	*Require “writing intensive” courses with multiple opportunities for feedback on writing.
<b>Develop leadership skills</b>	<ul style="list-style-type: none"> <li>* Descriptive data from student affairs on activity participation</li> <li>* Student focus groups</li> <li>* National Survey of Student Engagement (NSSE) results</li> </ul>	Student participation in campus governance is low.	<ul style="list-style-type: none"> <li>* Students work many hours off-campus for pay.</li> <li>* Students believe that the student government organization is ineffectual.</li> </ul>	<ul style="list-style-type: none"> <li>* Consider course, internship, or apprenticeship credit for student government participation.</li> <li>* Examine governance structure to revise, if necessary, in order to empower students.</li> </ul>
<b>Develop proficiency in academic and scholarly research skills</b>	<ul style="list-style-type: none"> <li>* User activity analysis of library resources</li> <li>* Library usage survey</li> </ul>	Analyses of user activity demonstrate overuse of non-academic databases in lieu of scholarly ones.	<ul style="list-style-type: none"> <li>* Students are not being taught the value of scholarly sources relative to popular sources.</li> <li>* Scholarly journals and other resources are not widely available in the library.</li> </ul>	<ul style="list-style-type: none"> <li>* Develop a required information literacy program that includes, among other things, examples of erroneous conclusions drawn from reviewing inappropriate sources.</li> <li>* Examine library budget with the goal of making more appropriate resources available.</li> <li>* Enter consortia that make scholarly materials easily obtainable for students.</li> </ul>
<b>Demonstrate career-skill preparation and preparedness for job-seeking</b>	<ul style="list-style-type: none"> <li>* Career Services appointment logs and attendance records for Career Services programs</li> <li>* Alumni surveys</li> </ul>	Very low numbers of students make use of career services programs.	<ul style="list-style-type: none"> <li>* Students are not aware of the services or their value.</li> <li>* Students who have used career services report low satisfaction.</li> </ul>	<ul style="list-style-type: none"> <li>* Require brief introduction to career services at first-year orientation and again in junior level general education courses; include testimonials from students who have used the services.</li> <li>* Revamp career services program to ensure better attainment of its goals.</li> </ul>



Figure 15

## Using Assessment Results at the Program Level: Preparing Students for Future Success

Learning Goal: Students will...	Direct and Indirect Measures	Outcome	Possible Reason or Hypothesis	Action Taken
<b>Be prepared for graduate and professional degree programs</b>	<ul style="list-style-type: none"> <li>* Departmental survey of graduating seniors and recent alumni</li> <li>* Data from benchmark institutions</li> </ul>	<ul style="list-style-type: none"> <li>* Student admittance rates to graduate and professional programs are low, compared to similar institutions' rates.</li> </ul>	<ul style="list-style-type: none"> <li>* Students are not being "coached" about the graduate school application process.</li> <li>* Students have not been exposed to experiences (e.g., undergraduate research) that enhance their chances of graduate school admissions.</li> </ul>	<ul style="list-style-type: none"> <li>* Enlist junior faculty members who have recently finished graduate school to develop a coaching program.</li> <li>* Incorporate a research, scholarship, or practicum requirement for students in a graduate or professional school "track."</li> </ul>
<b>Communicate competently in the major</b>	<ul style="list-style-type: none"> <li>* Scores on faculty-developed rubrics for final oral exam and final report in capstone course</li> </ul>	<ul style="list-style-type: none"> <li>* Student performance in capstone courses is poor, as measured by rubrics for oral presentations and written reports.</li> </ul>	<ul style="list-style-type: none"> <li>* Students are not receiving enough experience in communication in prerequisite major courses.</li> </ul>	<ul style="list-style-type: none"> <li>* Revamp departmental curriculum to require oral and written reports in every course.</li> <li>* Revamp syllabus of at least one required course to include multiple communication experiences.</li> </ul>
<b>Integrate competently knowledge and skills acquired in the major</b>	<ul style="list-style-type: none"> <li>* Departmental survey of graduating seniors</li> <li>* Grade distribution analysis of senior capstone course grades</li> </ul>	<ul style="list-style-type: none"> <li>* Survey results reveal that students think the capstone course is an "easy A."</li> <li>* Grade distribution reveals inflation.</li> </ul>	<ul style="list-style-type: none"> <li>* Capstone course is "watered-down" to account for a variety of previous course experience.</li> <li>* Capstone course does not demand true integration of previous learning.</li> </ul>	<ul style="list-style-type: none"> <li>* Change capstone course from a special-topics course to a course that requires an independent, integrative project.</li> <li>* Include a seminar component that makes students responsible for some of the course content.</li> </ul>

Learning Goal: Students will...	Direct and Indirect Measures	Outcome	Possible Reason or Hypothesis	Action Taken
<b>Exhibit fluency in a foreign language</b>	<ul style="list-style-type: none"> <li>* Scores on faculty-developed rubrics for oral presentation at end of intermediate-level courses</li> <li>* Standardized oral proficiency exam</li> </ul>	Students in a modern language program exhibit low levels of fluency in the language when tested after having taken several courses.	<ul style="list-style-type: none"> <li>* Courses are offered only two days a week.</li> <li>* Students have few opportunities for sustained practice in target language.</li> </ul>	<ul style="list-style-type: none"> <li>* Schedule courses for shorter periods, four or five days a week.</li> <li>* Introduce dedicated housing or separate floors for language students.</li> <li>* Place a fluent “graduate student in residence” in student housing.</li> <li>* Require a study abroad or immersion experience for language majors.</li> </ul>
<b>Demonstrate applied competency in the major</b>	Scores on faculty-developed rubrics	Students’ applied projects (e.g., design, engineering, fine arts) consistently reveal high levels of quality as evidenced by scores on rubrics designed to assess their efforts.	<ul style="list-style-type: none"> <li>* Students are very well prepared for applied projects in lower-level courses.</li> <li>* Could students be challenged even more?</li> </ul>	<ul style="list-style-type: none"> <li>* Require students to submit their work to external conferences or scholarly undergraduate journals.</li> <li>* Engage upper-class students as mentors for lower-class students.</li> </ul>
<b>Demonstrate competence in academic subject areas in the major</b>	Standardized disciplinary test	Standardized disciplinary tests (e.g., the ETS subject area tests, American Chemical Society Examination) reveal “spotty” or irregular command of the discipline.	<ul style="list-style-type: none"> <li>* Curriculum requirements uneven (i.e., no required courses in some target areas).</li> <li>* Courses poorly designed or delivered in target area (i.e., no permanent faculty member in that specialty).</li> </ul>	<ul style="list-style-type: none"> <li>* A team of faculty reassess course content.</li> <li>* Full-time faculty create syllabi for and mentor adjunct faculty.</li> <li>* Evaluate need for additional faculty time.</li> <li>* Provide stipends for excellent faculty to “retool” to teach in those content areas.</li> </ul>
<b>Demonstrate practical competence in the major</b>	<ul style="list-style-type: none"> <li>* Rating forms completed by practicum supervisor</li> <li>* Self-rating forms completed by students</li> </ul>	Students in nursing, education, or other programs requiring practicum experience are receiving lower than desirable scores from their supervisors.	<ul style="list-style-type: none"> <li>* Lower-level courses do not provide “mini-practicum” experiences.</li> <li>* Students have not been made aware of the importance of practica.</li> <li>* Little guidance is provided at practicum site.</li> </ul>	<ul style="list-style-type: none"> <li>* Interact with on-site practicum mentors to brainstorm about reason for the problem.</li> <li>* Revise prerequisite courses to include short practicum assignments.</li> </ul>

Figure 16

## Using Assessment Results at the Course Level: Ensuring Learning

[Note: Two of the goals are repeated on this chart to illustrate both “positive” and “negative” outcomes.]

Learning Goal: Students will...	Direct and Indirect Measures	Outcome	Possible Reason or Hypothesis	Action Taken
<b>Master course content</b>	<ul style="list-style-type: none"> <li>* Mid-term exam scores</li> <li>* Course evaluations</li> </ul>	Subject mastery is inadequate, as demonstrated by low in-class exam scores.	<ul style="list-style-type: none"> <li>* Students are not engaged with the subject matter.</li> <li>* Teaching format and exam format may not be compatible.</li> </ul>	<ul style="list-style-type: none"> <li>* Experiment with alternative teaching formats (e.g., problem-based learning, rather than lecture).</li> <li>* Create “test-blueprints” to acquaint students with expectations.</li> <li>* Analyze test format to determine if it is appropriate (e.g., Is an essay test being used when the material was presented as a series of facts requiring no analysis?) and change test format if warranted.</li> </ul>
<b>Master course content</b>	<ul style="list-style-type: none"> <li>* Mid-term exam scores</li> <li>* Course evaluations</li> </ul>	Subject mastery is very high, as demonstrated by especially high in-class exam scores.	<ul style="list-style-type: none"> <li>* Students are engaged with material.</li> <li>* Teaching format and exam format are compatible.</li> <li>* Students are aware of what competencies are necessary.</li> <li>* But are students being sufficiently challenged?</li> </ul>	<ul style="list-style-type: none"> <li>* Experiment with increasingly more difficult exams to gauge more accurately students’ potential.</li> <li>* Provide additional and especially challenging in-class or out-of-class assignments and assessment.</li> </ul>
<b>Exhibit discipline-specific writing skills</b>	<ul style="list-style-type: none"> <li>* Rubric-scored writing assignments</li> <li>* Course evaluations</li> </ul>	Average writing performance is especially high, as demonstrated by rubric-scored writing assignments.	<ul style="list-style-type: none"> <li>* Writing was effectively taught in prerequisite classes and/or is taught well in the target class.</li> <li>* But can students be challenged more, or can the course focus on the development of other skills?</li> </ul>	<ul style="list-style-type: none"> <li>* Engage students as mentors for other students who need help with writing.</li> <li>* Encourage students to submit their work to on- or off-campus literary or scholarly outlets for student work.</li> </ul>
<b>Exhibit discipline-specific writing skills</b>	<ul style="list-style-type: none"> <li>* Rubric-scored writing assignments</li> <li>* Course evaluations</li> </ul>	Average writing performance is unsatisfactory, as demonstrated by rubric-scored writing assignments.	<ul style="list-style-type: none"> <li>* Students were not prepared well in prerequisite classes.</li> <li>* Students were not receiving enough practice and feedback in target class.</li> </ul>	<ul style="list-style-type: none"> <li>* Require writing assignments with feedback in prerequisite classes.</li> <li>* Require writing assignments with feedback in target class.</li> <li>* Target specific aspects of student writing and focus assignments on them.</li> </ul>

Learning Goal: Students will...	Direct and Indirect Measures	Outcome	Possible Reason or Hypothesis	Action Taken
<b>Think critically and analytically</b>	<ul style="list-style-type: none"> <li>* Commercially-developed critical-thinking test</li> <li>* Problem-solving exercises</li> <li>* Applications problems</li> </ul>	Analytical skills/critical thinking are weak, as demonstrated by poor performance on problem-solving exercises or application problems.	<ul style="list-style-type: none"> <li>* Students have received little practice with these skills.</li> <li>* Faculty member has limited expertise in teaching these skills.</li> </ul>	<ul style="list-style-type: none"> <li>* Examine prerequisite courses for opportunities to engage in critical thinking, and revise appropriately.</li> <li>* Establish faculty-to-faculty mentor pairs, with each faculty member having expertise in an area the other lacks.</li> </ul>
<b>Demonstrate discipline-specific information literacy</b>	<ul style="list-style-type: none"> <li>* Rubric-scored term paper</li> <li>* Student survey or course evaluations</li> </ul>	Research skills are weak, as demonstrated by poor performance on research term paper.	<ul style="list-style-type: none"> <li>* Expectations for term paper were not clearly spelled out.</li> <li>* Appropriate search engines or sources were not available in the library.</li> <li>* Little practice with these skills on smaller assignments.</li> </ul>	<ul style="list-style-type: none"> <li>* Create a “blueprint” for the paper, clearly spelling out expectations.</li> <li>* Require an early non-graded first draft and provide feedback.</li> <li>* Examine library funding for possible reallocation.</li> <li>* Revamp required course to include several practice assignments drawing on scholarly research skills in the discipline.</li> </ul>
<b>Exhibit Oral Communication Skills</b>	Rubric-scored oral presentation	Oral presentation skills are weak, as demonstrated by rubric-scored presentation.	<ul style="list-style-type: none"> <li>* Prerequisite courses required little practice in oral presentation.</li> <li>* Students were not directly informed about the characteristics of a good presentation and/or the expectations for the assignment.</li> </ul>	<ul style="list-style-type: none"> <li>* Include an oral presentation component in early required courses with a comparable rubric for providing feedback.</li> <li>* Create teams of students to critique each others’ presentations before they are presented to the whole class, and include safeguards to make this a non-threatening, critically constructive situation.</li> </ul>
<b>Exhibit Quantitative Skills</b>	Faculty-developed mathematics test administered in an accounting class	Quantitative analysis skills are weak.	No appropriate prerequisite course is required.	Test students and assign them to a prerequisite course, if necessary.
<b>Exhibit High Student Morale and Satisfaction to Support Learning</b>	Teaching evaluations	There is dissatisfaction with course format, teaching style, and level of course difficulty, as demonstrated by teaching evaluations.	<ul style="list-style-type: none"> <li>* Over-dependence on lecture format</li> <li>* Little direct explanation of the value of the course or material.</li> <li>* Few opportunities for students to apply course content.</li> <li>* Faculty member inaccessible outside classroom.</li> </ul>	<ul style="list-style-type: none"> <li>* Experiment in small ways with non-lecture format (i.e., problem-based learning, group projects, “inter-teaching” in which students are expected to explain the material to each other).</li> <li>* Include an interactive discussion of the value of the material during the first class period, and regularly use applied examples to support course material.</li> <li>* Reconsider office hours schedule, or offer students opportunities to schedule appointments outside of office hours.</li> </ul>

## *A Message to Faculty Members:*

### **The Interconnectedness of Assessment, Teaching, and Learning**

*In the final section of this handbook we depart from the third person and move to the first and second person in order to speak directly to faculty whose lives are so connected with students in the classroom and in other learning situations on a daily basis.*

The staff of accrediting organizations or those who support accreditation by volunteering their services are very aware that an intense emphasis on student outcomes carries with it the danger of shifting the focus from teaching and learning to ill-considered data-collection. Therefore, the Middle States Commission on Higher Education advocates sincere attention to what faculty want students to learn; it does not advocate a “bean-counting” approach to assessment and improvement.

If you are convinced, on a theoretical level, that outcomes assessment is a well-intended and even a good thing, you may be thinking that you cannot possibly incorporate it into your classes, in your laboratories, or in your other pedagogical interactions with students. Yet, there are several strategies that can be accomplished without dramatic increases in faculty workload. Taken together, they can transform the way students learn.

The list presented in Figure 17 was developed from a collection of recent research and wisdom on when and how students learn best. Nine of the eleven items involve practices or characteristics that originate in individual classes. The two remaining characteristics are also directly related to the classroom. The list, for the most part, is self-explanatory, and some of these characteristics can be fostered with little change in a professor’s existing practices.

Here we elaborate on several of the items on the list, with an emphasis on how you can make small changes by engaging in assessment.

#### ➤ **Students learn effectively when they have opportunities to revise their work.**

If you currently do not grade multiple drafts of papers, projects, or lab reports, or provide multiple critiques of artwork or performances, consider taking at least one assignment and building in two or more reviews. On the earlier review or reviews, offer comments or grade with a rubric, but give students an opportunity to refine their work further before you assign a final grade. If necessary, consider making the project shorter or eliminating another graded assignment in order to provide time for the extra grading additional reviews will entail.

Another way to provide additional opportunities for students to revise their work is to initiate a system of peer review in which students share work with each other and review it according to pre-defined objective criteria.

#### ➤ **Students learn effectively when they understand course and program goals.**

Human beings take in information and learn new things much better when they have a framework upon which to rest new ideas. For instance, as an expert in your field, you can read a scholarly article much faster than can a novice. From prior experience, you know why the material presented in the article is important or not, whether it is controversial, whether it adds significantly to the current knowledge base in the area, and whether it appears to be a reflection of solid theory in your discipline. Without such background, the contents of the article would be meaningless or at least not as rich. If your syllabus has no goals or objectives listed, you are providing your students with no framework to help them understand where the course fits in with other courses, how the skills they will be acquiring translate to other

Figure 17

## Strategies to Improve Student Learning

There is increasing evidence that students learn most effectively when:

- They understand course and program goals and the characteristics of excellent work.
- They are academically challenged and encouraged to focus on developing higher-order thinking skills, such as critical thinking and problem solving, as well as discipline-specific knowledge.
- They spend more time actively involved in learning and less time listening to lectures.
- They engage in multidimensional “real world” tasks.
- Their learning styles are accommodated.
- They have positive interactions with faculty and work collaboratively with fellow students; all learners—students and professors—respect and value others as learners.
- They participate in out-of-class activities, such as co-curricular activities and service learning opportunities, that build on what they are learning in the classroom.
- Assignments and assessments are intertwined with learning activities and focus on the most important course and program goals.
- They have opportunities to revise their work.
- They reflect on what and how they have learned.
- They have a culminating “capstone” experience, such as a seminar, internship, independent study, research project, or thesis, that lets them synthesize what they have learned over the course of their college experience.

### Sources:

- Angelo, T. A. (1993, April). A “teacher’s dozen”: Fourteen general, research-based principles for improving higher learning in our classrooms. *AAHE Bulletin*, 3 (7), 13.
- Barr, R. B., & Tagg, J. (1995, Nov/Dec). From teaching to learning: A new paradigm for undergraduate education. *Change*, 27 (6), 12-25.
- Chickering, A. W., & Gamson, Z. (1987, July). Seven principles for good practice in undergraduate education. *AAHE Bulletin*, 39 (7), 5-10.
- Kuh, G. (2001, May/June). Assessing what really matters to student learning: Inside the National Survey of Student Engagement. *Change*, 33 (3), 10-17, 66.
- Mentkowski, M. & Associates. (2000). *Learning that lasts: Integrating learning, development, and performance in college and beyond*. San Francisco: Jossey-Bass.
- Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco: Jossey-Bass.
- What research says about improving undergraduate education. (1996, April). *AAHE Bulletin*, 48 (8), 5-8.

© 2002, Copyright by Linda Suskie. Reproduced with permission.

domains, or why they should be interested in the topic.

However, if objectives are listed, and they are made meaningful through a discussion of their importance and of the students' own objectives for the course, you will be modeling the behavior you use in your own scholarship. If you extend the discussion further and ask students to reflect on how class activities meet the objectives, you will be offering them a framework to evaluate all of their learning experiences, including those in other classes and in their future educational experiences.

➤ **Students learn most effectively when they (and you) reflect on what and how they have learned.**

If you have well-articulated goals for your course and if you have engaged students in a conversation about the importance of the course and its relationship to other aspects of their college program, the next step is to ensure that the pedagogical content of the course leads to achieving those goals. One technique that is not time consuming is to spend five minutes or so before each class listing the course goals or objectives for the class. Keep the record of your reflections for one semester, and at the end, look at it in aggregate. Are the goals supported relatively evenly? Is one goal over-represented (e.g., Is more time devoted to content than to process than seems appropriate given the goals)? Is there evidence that a large portion of time is spent on topics or activities that are not directly related to a course goal? Use the data you have collected to revise the course, perhaps only slowly over the course of several semesters, in order to help steer it toward the intended goals.

➤ **Students learn most effectively when assignments and assessments that are directly relevant to course goals are intertwined with learning activities and focus on the most important course and program goals.**

Consider an assignment that you give regularly or an exam format that you use most often. Does the assignment draw on or help teach the

particular capabilities that you are hoping to nurture in your students? Perhaps you assign a term paper each semester on a topic related to the subject area of your course. Is the sole purpose of the assignment to have students gain an in-depth knowledge of a particular topic? Do you have some tacit goals such as teaching students to engage in scholarly research, develop writing skills, or learn appropriate citation techniques—or is your goal simply to evaluate what they have been able to learn on their own? If these goals are tacit, rather than explicit, it is unlikely that your students will be aware of them, and thus unlikely that they will organize their work around them. If students are unable to connect the assignment to specific goals, the term paper is probably not an assessment of what they have learned in your class but, rather, an assessment of their “natural” competencies or characteristics, or prior learning.

In the case of examinations, many teaching faculty develop the habit of using the same format for exams over and over, without thinking of the congruence between what they want students to learn and the way they are assessing the learning. The format used for examinations may be the one that they found most congenial when they were students, the one that comes in a test bank with the text, or the one that is quickest to prepare or grade. However, every examination format has positive and negative characteristics.

For example, multiple-choice exams can be more psychometrically sound—more reliable and valid—if they are developed properly, than casually developed essay exams. Conversely, essay exams can sometimes allow for the assessment of a deeper, more synthetic (or analytic) exposition and certainly provide a venue for students to demonstrate their creativity or unique perspectives. Yet, no examination is valid if it is measuring something other than what it is intended to measure. Students frequently lament, “We didn’t know what you were going to put on the exam” or “You didn’t teach us the answers to the questions on the exam.” These students may be more sophisticated and less lazy than they appear. What students really might be saying is that this particular class did not prepare them to take this particular exam.

Some students, nevertheless, perform exceedingly well. These students would be a useful group to shed light on the connection between the class and the exam. Do they feel they could have answered some of the questions without ever having taken the course? This can happen in the case of essay questions that involve the use of logic or synthesis more than they do course content, or for which course content could have been gleaned outside of class. Do high-performing students feel that they could have done well on a multiple-choice test simply by memorizing the content of a text? An instructor may deliver brilliant and engaging lectures, but if he gives an exam on the content of the text instead, he has no assessment of what students learned during those great lectures.

A brief look at your assignments and exams relative to your goals and the content of your class presentations can be accomplished in a very short period of time. Greater insight can be gained by conducting some 10- or 15-minute “focus groups” with students about the congruence between the course and the assessments. Any changes that appear to be indicated can be made one at a time or over the course of a few semesters.

➤ **Students learn most effectively when they understand the characteristics of excellent work.**

The debate about grade inflation is complicated, although most professors probably identify easily with a specific perspective. The first common perspective is that if we would all just use a normal curve to grade, we wouldn't have grade inflation. The second perspective is that if we didn't tie teaching evaluations to faculty tenure and promotion, we wouldn't have grade inflation. The third perspective is that grade inflation is not necessarily a problem. If we are teaching for mastery—i.e. achievement of learning goals—then many students will earn high grades. Without taking sides, it is easy to make the case that grades mean very different things to the outside world, depending upon who assigns them. In a curve system, no one except the professor knows anything about students'

absolute performance. Some professors probably do grade generously hoping to receive good student evaluations. Even when professors' grades are meant to indicate mastery, we do not know which student mastered the material as a result of the course and which student had already mastered it before her arrival. It is clearly beyond the scope of a handbook on student learning assessment to solve the complex problem of incongruence among professors' grading systems. However, regardless of your perspective on grade inflation and within the context of your own class, you can ensure that grades are meaningful so that students recognize the characteristics of excellent work. After all, it is what the student learns, not the grade itself, that is most important.

Test blueprints are one way to make grades meaningful (see Chapter 3). The professor plots out what it is that a student is expected to master, and the professor might even create a hierarchy of mastery tied to various grades. After a student takes an examination, she can go back to the blueprint and reflect on her grade relative to the concepts she was expected to have mastered.

Rubrics (see Chapter 3) make the grading process more transparent, more accessible, and when well formulated, they are diagnostic for the student (Walvoord & Anderson, 1998). In reviewing his rubric scores, a student can pinpoint his areas of strength and weakness and develop his own learning goals to strengthen performance in specific areas.

Another accessible way to make grades more meaningful, although less structured than a rubric, is to write relevant and instructive comments on student work, including multiple-choice exams! Surprisingly, most students find few comments on their work that provide meaningful feedback, relegating tests and assignments to the category of summative assessment. It does take longer to write comments, but if they are important and strategically placed, there needn't be many.

How could one write meaningful comments on a multiple-choice exam? Questions on these exams fall into several categories, including memorization (Did the student acquire necessary



facts?); application (Given a scenario, can the student classify it, or extrapolate from it?); inference (If this happens, then what would happen?); and analogy (A is to b, as c is to d?). Students tend to err in patterns, frequently missing questions of the same type. Some guidance on which types of questions are causing difficulty for each student will almost certainly help him to focus his attention on the problem area next time.

As with the other areas discussed above, reflection on the meaning of grading—both the “score” and the comments offered—needn’t take a great deal of time, and modifications in one’s practices can be made gradually.

➤ **Students learn most effectively when their learning styles are accommodated.**

Recent pedagogical literature is replete with information about the varieties of learning styles that students can exhibit. Much of the work on learning styles has its genesis in traditional personality theory and its modern applied counterparts, such as the Myers-Briggs Type Indicator (Briggs-Myers, 1992), which classifies personality along four dimensions and yields sixteen types. Other heuristic impetus for work in the area of student learning styles comes from the theory of Kolb (1984) who defined four constituent processes involved in learning, Gardner (1983) who developed the now well-known concept of “multiple intelligences,” and Sternberg (1988) who posited three distinct kinds of intelligence. Appendix 8 offers some background information on learning styles in the form of Frequently Asked Questions (FAQs) for those who are interested in exploring this topic in greater depth.

The focus of the learning styles approach as it applies to teaching is that all students learn differently, either because of their personalities or their diverse patterns of cognitive ability, and that teaching formats within a class should be equally diverse to accommodate students’ learning needs. Because the most commonly used teaching method is the lecture-discussion format, some students with either personality types or learning profiles that prevent them from

assimilating easily information in this format may be learning less in the classroom than their counterparts who “prefer” such a format. This is an over-simplification of the learning styles approach, but the concept is not complicated. Different students learn best in different ways and in different settings, and failure to take this into account can impede learning in some students.

*For a variety of reasons, we are not advocating that professors assess each student’s learning profile and make associated individual accommodations (see Appendix 8).* However, any instructor can diversify his or her presentation and the types of experiences associated with a specific class to increase the chances of engaging students who may not thrive in a lecture-based classroom. Knowledge about learning styles can be used to modify other aspects of a course as well. Hartman (1995) suggests that putting students together in groups to work on a project can have the best outcomes when students’ learning styles are complementary, rather than similar. For instance, a group of students who all prefer theory or abstract thinking to concrete approaches may have difficulty organizing and getting started on the project, whereas another group of students, all of whom prefer detail and information-gathering over integration of ideas, may complete the project more efficiently, but may produce a less synthesized or mature product. Combining students with these styles will mix the best of both approaches.

**In parting...**

The purpose of defining goals and assessing learning is to improve learning through teaching. Teaching lies primarily in the hands of faculty members, and good learning cannot happen without their commitment and dedication. *Assessment, first and foremost, is a tool for faculty members to use as they do their very best to teach their students well.*

## Assessment Standards in *Characteristics of Excellence in Higher Education*

### Standard 7: Institutional Assessment

**The institution has developed and implemented an assessment process that evaluates its overall effectiveness in achieving its mission and goals and its compliance with accreditation standards.**

#### Context

Assessment may be characterized as the third element of a four-step planning-assessment cycle:

1. Developing clearly articulated written statements, expressed in observable terms, of key institutional and unit-level goals that are based on the involvement of the institutional community, as discussed under Standard 1 (Mission and Goals);
2. Designing intentional objectives or strategies to achieve those goals, as discussed under Standard 2 (Planning, Resource Allocation, and Institutional Renewal);
3. Assessing achievement of those key goals; and
4. Using the results of those assessments to improve programs and services, as discussed under Standard 2 (Planning, Resource Allocation, and Institutional Renewal), with appropriate links to the institution's ongoing planning and resource allocation processes.

The effectiveness of an institution rests upon the contribution that each of the institution's programs and services makes toward achieving the goals of the institution as a whole. This standard on institutional assessment thus builds upon all other accreditation standards, each of which includes periodic assessment of effectiveness as one of its fundamental elements. This standard ties together those assessments into an integrated whole to answer the question, "As an institutional community, how well are we collectively doing what we say we are doing?" and, in particular, "How do we support student learning, a

fundamental aspect of institutional effectiveness?" Because student learning is a fundamental component of the mission of most institutions of higher education, the assessment of student learning is an essential component of the assessment of institutional effectiveness and is the focus of Standard 14 (Assessment of Student Learning). Self-studies can thus document compliance with Standard 7 by summarizing the assessments of each accreditation standard into conclusions about the institution's overall achievement of its key goals.

The fundamental question asked in the accreditation process is, "Is the institution fulfilling its mission and achieving its goals?" This is precisely the question that assessment is designed to answer, making assessment essential to the accreditation process. Assessment processes help to ensure the following:

- Institutional and program-level goals are clear to the public, students, faculty, and staff;
- Institutional programs and resources are organized and coordinated to achieve institutional and program-level goals;
- The institution is indeed achieving its mission and goals; and
- The institution is using assessment results to improve student learning and otherwise advance the institution.

While the Commission expects institutions to assess institutional effectiveness, it does not prescribe a specific approach or methodology. The institution is responsible for determining its expected goals and the objectives or strategies for achieving them at each level (institutional and unit), assessment approaches and methodologies, sequence, and time frame. These may vary, based on the mission, goals, organization, and resources of the institution.

Whatever the approach, effective assessment processes are useful, cost-effective, reasonably accurate and truthful, carefully planned, and organized, systematic, and sustained.

**Useful** assessment processes help faculty and staff make appropriate decisions about improving programs and services, developing goals and plans, and making resource allocations. To assist with interpretation and use of assessment results, assessment measures and indicators have defined minimally acceptable performance targets. Because institutions, their students, and their environments are continually evolving, effective assessments cannot be static; they must be reviewed periodically and adapted in order to remain useful.

**Cost-effective** assessment processes yield dividends that justify the institution's investment in them, particularly in terms of faculty and staff time. To this end, institutions may begin by considering assessment measures, indicators, "flags," and "scorecards" already in place, such as retention, graduation, transfer, and placement rates, financial ratios, and surveys. New or refined measures may then be added for those goals and objectives for which evidence of achievement is not already available, concentrating on the institution's most important goals. Effective assessments are simple rather than elaborate and may focus on just a few key goals in each program, unit, and curriculum.

**Reasonably-accurate and truthful** assessment processes yield results that can be used with confidence to make appropriate decisions. Because there is no one perfectly accurate assessment tool or strategy, institutions should use multiple kinds of measures to assess goal achievement. Assessments may be quantitative and/or qualitative and developed locally or by an external organization. All assessment tools and strategies should clearly relate to the goals they are assessing and should be developed with care; they should not be merely anecdotal information nor collections of information that happen to be on hand. Strategies to assess student learning should include direct—clear, visible, and convincing—evidence, rather than solely indirect evidence of student learning such as surveys and focus groups.

**Planned** assessment processes that purposefully correspond to institutional goals that they are intended to assess promote attention to those goals and ensure that disappointing outcomes are appropriately addressed. Institutions often have a variety of plans, such as a strategic plan, academic plan, financial plan, enrollment plan, capital facilities master plan, and technology plan. Just as such plans should be interrelated to ensure that they work synergistically to advance the institution, assessments should also be interrelated. At many institutions, effective institutional planning begins with academic planning, which in turn drives the other plans. If the academic plan calls for a new academic program, for example, the technology plan should ensure faculty and students in the new program will be able to use appropriate instructional technologies. Assessments of the technology plan should evaluate not just whether instructional technologies have been put in place but also how effectively those technologies have helped students to achieve the program's key learning outcomes.

**Organized, systematized, and sustained** assessment processes are ongoing, not once-and-done. There should be clear interrelationships among institutional goals, program- and unit-level goals, and course-level goals. Assessments should relate clearly to important goals, and improvements should clearly stem from assessment results.

As noted earlier, because student learning is a fundamental component of the mission of most institutions of higher education, the assessment of student learning is an essential component of the assessment of institutional effectiveness. An institution may therefore create institutional effectiveness documentation that includes a component on assessing student learning (see Standard 14: Assessment of Student Learning), or it may create a bridge between two separate sets of documentation, one for the assessment of student learning and one for other aspects of institutional effectiveness.

A commitment to the assessment of institutional effectiveness requires a parallel commitment to ensuring its use. Assessment information, derived in a manner appropriate to the institution and to its desired outcomes, should be available to and used by those who develop institutional goals and carry out strategies to achieve them. As discussed under Standard 2 (Planning, Resource Allocation, and Institutional Renewal), an accredited institution uses the results of assessment for institutional renewal: to maintain, support, and improve its programs and services. Assessment information should be used as a basis for assessing the institution's effectiveness in achieving its stated goals, for monitoring and improving the environment for student learning, and for enhancing overall student success; to these ends, it should be linked to the institution's ongoing planning and resource allocation processes.

Assessment results also should be used to evaluate the assessment process itself, leading to modifications that improve its relevance and effectiveness.

## Fundamental Elements of Institutional Assessment

An accredited institution is expected to possess or demonstrate the following attributes or activities:

- documented, organized, and sustained assessment process to evaluate and improve the total range of programs and services; achievement of institutional mission, goals, and plans; and compliance with accreditation standards that meets the following criteria:
  - ✧ a foundation in the institution's mission and clearly articulated institutional, unit-level, and program-level goals that encompass all programs, services, and initiatives and are appropriately integrated with one another (see Standards 1: Mission and Goals and 2: Planning, Resource Allocation, and Institutional Renewal);
  - ✧ systematic, sustained, and thorough use of multiple qualitative and/or quantitative measures that:

- ✧ maximize the use of existing data and information;
- ✧ clearly and purposefully relate to the goals they are assessing;
- ✧ are of sufficient quality that results can be used with confidence to inform decisions;
- ✧ support and collaboration of faculty and administration;
- ✧ clear realistic guidelines and a timetable, supported by appropriate investment of institutional resources;
- ✧ sufficient simplicity, practicality, detail, and ownership to be sustainable;
- ✧ periodic evaluation of the effectiveness and comprehensiveness of the institution's assessment process;
- evidence that assessment results are shared and discussed with appropriate constituents and used in institutional planning, resource allocation, and renewal (see Standard 2: Planning, Resource Allocation, and Institutional Renewal) to improve and gain efficiencies in programs, services and processes, including activities specific to the institution's mission (e.g., service, outreach, research); and
- written institutional (strategic) plan(s) that reflect(s) consideration of assessment results.

Institutions and evaluators must consider the totality that is created by the fundamental elements and any other relevant institutional information or analysis. Fundamental elements and contextual statements should not be applied separately as checklists. Where an institution does not possess or demonstrate evidence of a particular Fundamental Element, the institution may demonstrate through alternative information and analysis that it meets the standard.

## Optional Analysis and Evidence

In addition to the evidence inherent within or necessary to document the fundamental elements above, the following, although not required, may facilitate the institution's own analysis relative to this accreditation standard:

- analysis of the institutional culture for assessing institutional effectiveness, including:
  - ✧ the views of faculty and administrators on assessment;
  - ✧ faculty and administrators' understanding of their roles in assessing institutional effectiveness;
  - ✧ campus-wide efforts to encourage, recognize, and value efforts to assess institutional effectiveness and to improve programs and services;
- analysis of the quality and usefulness of institutional support for assessment efforts, including the quality and usefulness of:
  - ✧ written statements of expectations for assessment work;
  - ✧ policies and governance structures to support institutional assessment;
  - ✧ administrative, technical, and financial support for institutional assessment activities;
  - ✧ professional development opportunities and resources for faculty and staff to learn how to assess institutional effectiveness and how to use the results;
- clear, appropriate criteria for determining whether key institutional goals and objectives have been achieved;
- analysis of whether the institution has sufficient, convincing, written evidence that it is achieving its mission and its key institutional goals;
- analysis of results of surveys of students and other relevant groups;
- review of evaluations of special, mission driven programs or projects, with

recommendations for improvement, and evidence of action based on recommendations;

- evidence that institutional assessment findings are used to:
  - ✧ improve student success;
  - ✧ review and improve programs and services;
  - ✧ plan, conduct, and support professional development activities;
  - ✧ assist in planning and budgeting for the provision of programs and services;
  - ✧ support decisions about strategic goals, plans, and resource allocation;
  - ✧ inform appropriate constituents about the institution and its programs;
- evidence of renewal strategies, made in response to assessment results [included also under Standard 2 Optional Analyses]; or
- analysis of evidence that renewal strategies made in response to assessment results have had the desired effect in improving programs, services, and initiatives.

## Standard 14: Assessment of Student Learning

**Assessment of student learning demonstrates that, at graduation, or other appropriate points, the institution’s students have knowledge, skills, and competencies consistent with institutional and appropriate higher education goals.**

### Context

Assessment of student learning may be characterized as the third element of a four-step teaching-learning-assessment cycle:

1. Developing clearly articulated written statements, expressed in observable terms, of key learning outcomes: the knowledge, skills, and competencies that students are expected to exhibit upon successful completion of a course, academic program, co-curricular program, general education requirement, or other specific set of experiences, as discussed under Standard 11 (Educational Offerings);
2. Designing courses, programs, and experiences that provide intentional opportunities for students to achieve those learning outcomes, again as discussed under Standard 11;
3. Assessing student achievement of those key learning outcomes; and
4. Using the results of those assessments to improve teaching and learning.

This standard on assessment of student learning builds upon Standards 11 (Educational Offerings), 12 (General Education), and 13 (Related Educational Offerings), each of which includes assessment of student learning among its fundamental elements. This standard ties together those assessments into an integrated whole to answer the question, “Are our students learning what we want them to learn?” Self-studies can thus document compliance with Standard 14 by summarizing the assessments of Standards 11 through 13 into conclusions about overall achievement of the institution’s key student learning outcomes.

**Because student learning is at the heart of the mission of most institutions of higher education, the assessment of student learning is an essential component of the assessment of institutional effectiveness (see Standard 7: Institutional Assessment), which additionally monitors the environment provided for teaching and learning and the achievement of other aspects of the institution’s mission, vision, and strategic goals and plans.**

The fundamental question asked in the accreditation process is, “Is the institution fulfilling its mission and achieving its goals?” This is precisely the question that assessment is designed to answer, making assessment essential to the accreditation process. Assessment processes help to ensure the following:

- Institutional and program-level goals are clear to the public, students, faculty, and staff;
- Institutional programs and resources are organized and coordinated to achieve institutional and program-level goals;
- The institution is providing academic opportunities of quality;
- The institution is indeed achieving its mission and goals; and
- Assessment results help the institution to improve student learning and otherwise advance the institution.

Assessment is not an event but a process that is an integral part of the life of the institution, and an institution should be able to provide evidence that the assessment of student learning outcomes and use of results is an ongoing institutional activity. While some of the impact of an institution on its students may not be easily or immediately measured—some institutions, for example, aim for students to develop lifelong habits that may not be fully developed for many years—the overall assessment of student learning is expected whatever the nature of the institution, its mission, the types of programs it offers, or the manner in which its educational programs are delivered and student learning facilitated.

While the Commission expects institutions to assess student learning, it does not prescribe a specific approach or methodology. The institution is responsible for determining its expected learning outcomes and strategies for achieving them at each level (institutional, program, and course), assessment approaches and methodologies, sequence, and time frame. These may vary, based on the mission, goals, organization, and resources of the institution. Whatever the approach, effective assessment processes are useful, cost-effective, reasonably accurate and truthful, carefully planned, and organized, systematic, and sustained.

**Useful** assessment processes help faculty and staff make appropriate decisions about improving programs and services, developing goals and plans, and making resource allocations. To assist with interpretation and use of assessment results, assessment measures and indicators have defined minimally acceptable performance targets. Because institutions, their students, and their environments are continually evolving, effective assessments cannot be static; they must be reviewed periodically and adapted in order to remain useful.

**Cost-effective** assessment processes are designed so that their value is in proportion to the time and resources devoted to them. To this end, institutions can begin by considering assessment measures already in place, including direct evidence such as capstone projects, field experience evaluations, and performance on licensure examinations and indirect evidence such as retention and graduation rates and alumni surveys. New or refined measures can then be added for those learning outcomes for which direct evidence of student learning is not already available, concentrating on the most important institutional and program-level learning outcomes. Effective assessments are simple rather than elaborate and may focus on just a few key goals in each program, unit, and curriculum.

**Reasonably-accurate and truthful** assessment processes yield results that can be used with confidence to make appropriate decisions. Such assessment processes have the following characteristics:

- ✧ Because there is no one perfectly accurate assessment tool or strategy, institutions should use multiple kinds of measures to assess goal achievement. Assessments may be quantitative and/or qualitative and developed locally or by an external organization.
- ✧ Assessment tools and strategies should be developed with care; they should not be not merely anecdotal information nor collections of information that happen to be on hand.
- ✧ Student learning assessment processes should yield direct—clear, visible, and convincing—evidence of student learning. Tangible examples of student learning, such as completed tests, assignments, projects, portfolios, licensure examinations, and field experience evaluations, are direct evidence of student learning. Indirect evidence, including retention, graduation, and placement rates and surveys of students and alumni, can be vital to understanding the teaching-learning process and student success (or lack thereof), but such information alone is insufficient evidence of student learning unless accompanied by direct evidence. Grades alone are indirect evidence, as a skeptic might claim that high grades are solely the result of lax standards. But the assignments and evaluations that form the basis for grades can be direct evidence if they are accompanied by clear evaluation criteria that have a demonstrable relationship to key learning goals.

**Planned** assessment processes that clearly and purposefully correspond to learning outcomes that they are intended to assess promote attention to those goals and ensure that disappointing outcomes are appropriately addressed.

**Organized, systematized, and sustained** assessment processes are ongoing, not once-and-done. There should be clear interrelationships among institutional goals, program- and unit-level goals, and course-level

goals. Assessments should clearly relate to important goals, and improvements should clearly stem from assessment results.

As noted earlier, because student learning is a fundamental component of the mission of most institutions of higher education, the assessment of student learning is an essential component of the assessment of institutional effectiveness. An institution may therefore create institutional effectiveness documentation that includes a component on assessing student learning (see Standard 14: Assessment of Student Learning), or it may create a bridge between two separate sets of documentation, one for the assessment of student learning and one for other aspects of institutional effectiveness.

The improvement of overall educational quality and the enhancement of effective teaching and learning is most likely to occur when faculty and administrators work together to implement a sound, institution-wide program of assessment. Because the faculty guide decisions about curriculum and pedagogy, the effective assessment of student learning is similarly guided by the faculty and supported by the administration.

A commitment to assessment of student learning requires a parallel commitment to ensuring its use. Assessment information, derived in a manner appropriate to the institution and its desired academic outcomes, should be available to and used by those who develop and carry out strategies that will improve teaching and learning.

Assessment results should also be used to evaluate the assessment process itself, leading to modifications that improve its relevance and effectiveness.

## **Fundamental Elements of Assessment of Student Learning**

An accredited institution is expected to possess or demonstrate the following attributes or activities.

- clearly articulated statements of expected student learning outcomes (see Standard 11: Educational Offerings), at all levels (institution, degree/program, course) and for

all programs that aim to foster student learning and development, that are:

- ✧ appropriately integrated with one another;
- ✧ consonant with the institution's mission; and
- ✧ consonant with the standards of higher education and of the relevant disciplines;
- a documented, organized, and sustained assessment process to evaluate and improve student learning that meets the following criteria:
  - ✧ systematic, sustained, and thorough use of multiple qualitative and/or quantitative measures that:
    - ◆ maximize the use of existing data and information;
    - ◆ clearly and purposefully relate to the goals they are assessing;
    - ◆ are of sufficient quality that results can be used with confidence to inform decisions; and
    - ◆ include direct evidence of student learning;
  - ✧ support and collaboration of faculty and administration;
  - ✧ clear, realistic guidelines and timetable, supported by appropriate investment of institutional resources;
  - ✧ sufficient simplicity, practicality, detail, and ownership to be sustainable; and
  - ✧ periodic evaluation of the effectiveness and comprehensiveness of the institution's student learning assessment processes;
- assessment results that provide sufficient, convincing evidence that students are achieving key institutional and program learning outcomes;
- evidence that student learning assessment information is shared and discussed with appropriate constituents and is used to improve teaching and learning; and



- documented use of student learning assessment information as part of institutional assessment.

Institutions and evaluators must consider the totality that is created by the fundamental elements and any other relevant institutional information or analysis. Fundamental elements and contextual statements should not be applied separately as checklists. Where an institution does not possess or demonstrate evidence of a particular Fundamental Element, the institution may demonstrate through alternative information and analysis that it meets the standard.

## Optional Analysis and Evidence

In addition to the evidence inherent within or necessary to document the fundamental elements above, the following, although not required, may facilitate the institution's own analysis relative to this accreditation standard:

- analysis of institutional support for student learning assessment efforts, including:
  - ✧ written statements of expectations for student learning assessment work;
  - ✧ policies and governance structures to support student learning assessment;
  - ✧ administrative, technical, and financial support for student learning assessment activities and for implementing changes resulting from assessment; and
  - ✧ professional development opportunities and resources for faculty to learn how to assess student learning, how to improve their curricula, and how to improve their teaching;
- analysis of the clarity and appropriateness of standards for determining whether key learning outcomes have been achieved;
- evidence of workable, regularized, collaborative institutional processes and protocols for ensuring the dissemination, analysis, discussion, and use of assessment results among all relevant constituents within a reasonable schedule;
- analysis of the use of student learning assessment findings to:
  - ✧ assist students in improving their learning;
  - ✧ improve pedagogies, curricula and instructional activities;
  - ✧ review and revise academic programs and support services;
  - ✧ plan, conduct, and support professional development activities;
  - ✧ assist in planning and budgeting for the provision of academic programs and services;
  - ✧ support other institutional assessment efforts (see Standard 7: Institutional Assessment) and decisions about strategic goals, plans, and resource allocation; and
  - ✧ inform appropriate constituents about the institution and its programs;
- analysis of evidence that improvements in teaching, curricula, and support made in response to assessment results have had the desired effect in improving teaching, learning, and the success of other activities;
- analysis of the institutional culture for assessing student learning, including:
  - ✧ the views of faculty and institutional leaders on assessment;
  - ✧ faculty members' understanding of their roles in assessing student learning;
  - ✧ the quality and usefulness of institutional support for student learning assessment efforts;
  - ✧ campus-wide efforts to encourage, recognize, and value efforts to assess student learning and to improve curricula and teaching;
  - ✧ evidence of collaboration in the development of statements of expected student learning and assessment strategies;
- evidence that information appropriate to the review of student retention, persistence, and attrition, is used to reflect whether these are consistent with student and institutional

expectations [also included in Standard 8 Optional Analyses];

- evidence of the utilization of attrition information to ascertain characteristics of students who withdraw prior to attaining their educational objectives and, as appropriate, implementation of strategies to improve retention [also included under Optional Analyses in Standard 8];
- analysis of teaching evaluations, including identification of good practices; or
- analysis of course, department or school reports on classroom assessment practices and their outcomes, including grading approaches and consistency.

## Appendix 2

# Enhancing the Campus Climate for Assessment: Questions for Academic Leaders

### **What is your personal commitment to assessment?**

- Are you sufficiently familiar with current thinking about the principles and practice of assessment?
- Are you comfortable with the concept of assessment? Have you worked through any reservations you have about assessment?
- Do you understand why assessment is important?
- Are you personally committed to sharing leadership of assessment with the faculty?

### **How do you stimulate interest in assessment?**

- Do you promote assessment when you talk formally and informally with faculty, students, and staff?
- Do you sponsor consultants, speakers, and forums on assessment? Do you support these programs with your active presence?
- Do you explain to faculty, students, and staff how assessment findings affect major decisions that you and your colleagues make?
- Do you have communication channels with your campus assessment committee(s)? Do you actively use them to promote assessment?

### **How do you help provide the people who will help the campus focus on assessment?**

- Do you see faculty vacancies as an opportunity to move substantively toward promoting a learning-centered environment?
- Do you give hiring preference to faculty applicants who have documented success in creating a learning-centered environment for their students and in using assessment to strengthen teaching and learning?
- Do you ask faculty and staff applicants to demonstrate their skills in promoting active learning and their skills in assessment?

### **How do you give faculty incentives to focus on assessment?**

- Do you offer ample incentives for faculty and staff (e.g., promotion/tenure/merit considerations, reassigned time, budget supplements) to refocus their work in ways that promote a learning-centered environment and/or strengthen assessment?
- Are you promoting a learning-centered environment and strengthening assessment major goals for your institution?
- Are you promoting a learning-centered environment and strengthening assessment major goals for you personally?
- Do you require proposals for new programs to include plans for assessing student learning?

Continued on next page ã

### **How do you provide the training to enable faculty to strengthen assessment?**

- Do you encourage your faculty and staff to participate in campus and off-campus professional development programs on assessment?
- Do you alert your campus's teaching/learning center and/or assessment officer to faculty and staff needs for professional development on assessment?
- Do you fund faculty and staff travel to assessment conferences, institutes, and workshops?

### **How do you provide the resources to enable faculty to strengthen assessment?**

- Do you give any special funding to programs that make the most progress in strengthening and using assessment?
- Do you provide both "seed money" and sustained or other special funding for initiatives that significantly strengthen assessment?
- Do you encourage your institution to give priority to fundraising for programs and activities that make assessment a priority?
- Do you encourage and honor faculty who seek grants for resources that will promote a learning-centered environment and strengthen assessment?

### **How do you help the faculty focus their time on assessment?**

- Do you make assessment a focus of program reviews?
- Do you encourage departments to set department goals that contribute substantively toward promoting a learning-centered environment and strengthening assessment?
- Do you encourage and reward scholarship of teaching as a scholarly activity?
- Do you help faculty find the time for assessment initiatives by helping to minimize paperwork and by relieving them of less-critical responsibilities?

### **How do you encourage measurable outcomes of assessment endeavors?**

- Do you track the number of programs that make major progress in strengthening assessment?
- Do you track the percent of courses/sections that use the most appropriate assessment tools and strategies?
- Do you track the percent of students who participate in the embedded assessments of higher-order thinking skills?
- Do you track resource utilization to see how well it supports assessment?
- Do you develop other key performance indicators for assessment?

### **How do you celebrate and reward assessment achievements?**

- Do you announce noteworthy student and faculty accomplishments to internal and external constituents?
- Do you create celebrations of assessment achievements, consistent with campus culture?
- Do you provide special resources (e.g., revenue sharing) to those making extraordinary contributions to assessment?

## Appendix 3

# Assessment Practices Quiz

Mark a T next to those statements that accurately describe the Middle States Commission's views on assessment of student learning. Mark an F next to those statements that do **not** accurately describe the Commission's views on assessment of student learning.

1. \_\_\_\_ Published tests are always preferable to locally developed assessment measures.
2. \_\_\_\_ Class assignments can be used to assess the learning goals of academic programs.
3. \_\_\_\_ Tests with numeric scores are preferable to qualitative measures such as focus groups.
4. \_\_\_\_ Surveys of student satisfaction with a program are insufficient evidence of what students have learned.
5. \_\_\_\_ Every learning outcome of every course and program must be assessed.
6. \_\_\_\_ All students should be assessed to demonstrate the effectiveness of a course or program; a sample of students is inadequate.
7. \_\_\_\_ Goals should not be changed after they are selected.
8. \_\_\_\_ The same assessment measures should be used in every assessment cycle.
9. \_\_\_\_ Grades alone are not direct evidence of student learning.
10. \_\_\_\_ The primary purposes of assessment are to maintain accreditation and satisfy external stakeholders; therefore, it is appropriate to schedule assessment cycles so that they coincide with accreditation self-studies.
11. \_\_\_\_ The most effective way to create an assessment plan is to adopt the assessment plan of another institution.
12. \_\_\_\_ Assessment efforts should focus on what is learned in academic courses and programs; assessing what is learned in out-of-class activities is not important.
13. \_\_\_\_ Self-report measures can yield useful information about student learning.
14. \_\_\_\_ The assessment of educational effectiveness and the assessment of institutional effectiveness are not related.

## Appendix 4

# Key to “Assessment Practices Quiz”

1. **False** Published tests are always preferable to locally developed assessment measures. [*Both published and locally developed instruments have pros and cons, and both may have a place in an assessment program.*]
2. **True** Class assignments can be used to assess the learning goals of academic programs. [*Class assignments, especially in senior capstone courses, can be valuable sources of “embedded” information on how well students are achieving the major goals of a program.*]
3. **False** Tests with numeric scores are preferable to qualitative measures such as focus groups. [*Qualitative and quantitative measures offer different perspectives to an assessment program, and both can be valuable.*]
4. **True** Surveys of student satisfaction with a program are insufficient evidence of what students have learned. [*Student satisfaction surveys are not direct measures of student learning. They can, however, be informative indirect measures of learning.*]
5. **False** Every learning outcome of every course and program must be assessed. [*Only the key learning outcomes of courses and programs need be assessed on a regular basis.*]
6. **False** All students should be assessed to demonstrate the effectiveness of a course or program; a sample of students is inadequate. [*Samples can be cost-effective sources of information, provided that the samples are representative of all students, and the samples are sufficiently large so that the results can be generalized.*]
7. **False** Goals should not be changed after they are selected. [*Goals should be modified whenever it becomes clear that revising them would improve the student learning experience.*]
8. **False** The same assessment measures should be used in every assessment cycle. [*Assessment strategies can be implemented on a staggered basis and can be modified whenever it is clear that a new or revised strategy would be more useful.*]
9. **True** Grades alone are not direct evidence of student learning. [*Grades alone do not tell us exactly what a student has and has not learned. The information upon which grades are based—tests, student papers and projects, and the like—are direct evidence of student learning.*]
10. **False** The primary purposes of assessment are to maintain accreditation and satisfy external stakeholders; therefore it is appropriate to schedule assessment cycles so that they coincide with accreditation self-studies. [*The primary purpose of assessment is to improve student learning. Assessment should be systematic, continuous, and ongoing.*]
11. **False** The most effective way to create an assessment plan quickly is to adopt the assessment plan of another institution. [*Although an institution may choose to adapt some of the features of another institution’s assessment plan, each institution should develop an assessment plan that is tailored to its own culture, mission, and needs.*]
12. **False** Assessment efforts should focus on what is learned in academic courses and programs; assessing what is learned in out-of-class activities is not important. [*Both in-class and out-of-class activities include valuable learning opportunities that should be assessed.*]
13. **True** Self-report measures can yield useful information about student learning. [*Asking students to reflect upon their learning experiences can yield valuable insights into what they have and have not learned, especially their attitudes and values.*]
14. **False** The assessment of educational effectiveness and the assessment of institutional effectiveness are not related. [*Because teaching and learning are fundamental missions of every institution of higher education, the assessment of educational effectiveness is a major component of the assessment of institutional effectiveness.*]

## Appendix 5

# Department/Program Student Outcomes Survey

1. Does your department require a Capstone/Senior Culminating Experience?  Yes;  No

If yes, what form does this experience take?

A choice among several options (*Check all options available to students*):

- |   |  |  |
|---|--|--|
| <input type="checkbox"/> Senior Thesis    | <input type="checkbox"/> Topical Seminar         | <input type="checkbox"/> Independent Study |
| <input type="checkbox"/> Honors Thesis    | <input type="checkbox"/> Service Learning Course | <input type="checkbox"/> Other             |
| <input type="checkbox"/> Research Seminar | <input type="checkbox"/> Internship              |  |

A single course or requirement that must be completed by all students  
(*Check which form this requirement takes*):

- |   |  |  |
|---|--|--|
| <input type="checkbox"/> Senior Thesis    | <input type="checkbox"/> Topical Seminar         | <input type="checkbox"/> Independent Study |
| <input type="checkbox"/> Honors Thesis    | <input type="checkbox"/> Service Learning Course | <input type="checkbox"/> Student Teaching  |
| <input type="checkbox"/> Research Seminar | <input type="checkbox"/> Internship              | <input type="checkbox"/> Other             |

2. Is this a Capstone/Culminating Experience required by disciplinary accreditation or for professional certification?  Yes;  No

3. Please give examples of how your department adapts its curricula as a result of student performance in capstone experiences.

4. Does your department administer its own surveys or questionnaires to current students?  Yes;  No

If yes, when are these questionnaires or surveys administered? (*Check all that apply.*)

- First Year       Sophomore Year       Junior Year       Senior Year

Which of the following kinds of information is gathered on these surveys? (*Check all that apply.*)

- |   |   |  |
|---|---|--|
| <input type="checkbox"/> Graduate school applications | <input type="checkbox"/> Jobs offered                 | <input type="checkbox"/> Leadership activities                                       |
| <input type="checkbox"/> Graduate school acceptances  | <input type="checkbox"/> Jobs accepted                | <input type="checkbox"/> Satisfaction with the institution                           |
| <input type="checkbox"/> Graduate school chosen       | <input type="checkbox"/> Expected salary              | <input type="checkbox"/> Satisfaction with department                                |
| <input type="checkbox"/> Career plans                 | <input type="checkbox"/> Salary for accepted job      | <input type="checkbox"/> Satisfaction with major                                     |
| <input type="checkbox"/> Jobs applied for             | <input type="checkbox"/> Community service activities | <input type="checkbox"/> Satisfaction with teaching                                  |
|   |   | <input type="checkbox"/> Perceptions of how their education could have been improved |
|   |   | <input type="checkbox"/> Other   |

5. Please give examples of how your department adapts its curricula as a result of student surveys or questionnaires.





13. Can you give an examples of how information about scores on professional licensure examinations, professional certification examinations, or GREs are used to adapt your curriculum?  
Please add an addendum if necessary.
14. Does your department collect student portfolios?  Yes;  No  
If yes, are they collected from:  All students  A random sample  
If your department uses portfolios, briefly describe the types of material included in them.  
Please add an addendum if necessary.
15. Can you give examples of how your department adapts its curricula as a result of information from student portfolios?
16. Does your department keep an ongoing record of student accomplishments (e.g., student presentations at conferences, student gallery showings of artwork, student publications, student productions, etc.)?  Yes;  No
17. Does your department keep an ongoing record of student/faculty collaborative research and scholarship (e.g., presentations, publications, etc.)?  Yes;  No
18. Does your department require students to present their work to an audience of their peers and/or faculty?  Yes;  No  
If yes, list the course(s) for which such a presentation is required.
19. Does your department keep a record of competitive scholarships, fellowships, internships, or grants awarded to or won by your students?  Yes;  No
20. Does your department assess achievement of the general education goals within your program?  Yes;  No  
If yes, list the goals you assess, and briefly describe how they are assessed.  
Please add an addendum if necessary.
21. Can you identify other information-gathering techniques, not listed in this survey, that you use to assess student learning at the department or program level? Please add an addendum if necessary.
22. Which techniques for assessing student learning at the program or departmental level do your faculty find most useful and informative? Please add an addendum if necessary.

Source: The College of New Jersey, Ewing, NJ. Reproduced and adapted with permission.

Appendix 6

## Learning Goals and Assessment Techniques

This worksheet can help faculty and staff begin to define important learning goals and to determine appropriate assessment techniques.

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>
	<b>Important Goals: What students can do after completing the program (course, activity)</b>	<b>How do students learn to do this?</b>	<b>What information or evidence is there that students are learning this?</b>	<b>How has this information been used to help students learn?</b>	<b>What additional evidence is needed to understand how well students are learning this?</b>	<b>What possible new or improved assessment tools or techniques might be used?</b>
<b>1.</b>						
<b>2.</b>						
<b>3.</b>						

## Appendix 7

# From Effect to Cause: A Brainstorming Exercise

The scenarios presented here can be used as the focus of a brainstorming exercise to help faculty and staff members get a sense of the kinds of changes in curriculum and practice that may result from assessment.

**Scenario 1.** Faculty in the Biophysics Department agree that student majors should be able to make effective oral presentations of their research findings, but they are not satisfied with the quality of the oral presentations made by their seniors. Unfortunately, they can't find a place in the curriculum for students to practice preparing and making oral presentations. All the faculty agree that they have so much content to cover in their courses that they don't have time to teach students how to make effective oral presentations and then listen to them. How might the faculty address this?

**Scenario 2.** Senior Biology majors at Roselyn College scored poorly on the botany section of the XYZ National Biology Test. Some faculty believe that this is not a concern, because virtually all Roselyn Biology graduates go on to careers in the health and medical fields. Others believe that a grounding in botany is essential to being a well-rounded biologist. How might the faculty resolve this?

**Scenario 3.** In blind reviews, 85% of Cultural Anthropology senior theses were scored "outstanding" in terms of clarity, organization, the comprehensiveness of their review of scholarly literature, and the soundness of their analysis and conclusions. Five percent were scored "very good," 5% "adequate," and 5% "inadequate." How might the faculty use this information?

**Scenario 4.** The faculty members of the European Studies Department agree that their student majors should be able to summarize the principles or teachings of the major ancient Greek philosophers. Unfortunately, a review of a sample of student papers shows that the students are generally poor at doing this. To make matters worse, there is only one course in the department that covers ancient Greek philosophy, taught by a senior faculty member who adamantly refuses to consider modifying what or how he teaches. What might the department do?

**Scenario 5.** One of the goals of the Organizational Leadership program is that students are able to "write clearly and effectively." Although Organizational Leadership majors are asked to write term papers in at least six department courses, their writing quality is nonetheless inadequate by the time they become seniors. Faculty are quick to point to the woefully poor writing skills of entering freshmen and equally quick to blame the English Department for not bringing students' writing skills up to par in freshman composition classes. What, if anything, might be done to improve students' writing skills before they graduate?

## Student Learning Styles: Frequently Asked Questions

### What is “learning style”?

There is no one universally-accepted definition of the term learning style, but the most frequently cited definition appears to be cognitive, affective, and physiological factors that affect how learners perceive, interact with, and respond to the learning environment (Keefe, 1979).

### What is the predominant conception of learning style?

There are many models and instruments for categorizing learning styles, but they have not yet been integrated into an overall learning style theory (Bonham, 1988a; Rayner & Riding, 1997). As Vincent and Ross (2001) note, “Professional educators...are unable to form a consensus regarding the establishment of a single set of accepted principles.” Instruments have been developed from at least six models:

1. Field dependence/field independence (Group Embedded Figures Test)
2. Jungian models (Myers-Briggs Type Indicator, Gregorc Style Delineator, Keirsev Temperament Sorter II, Kolb Learning Style Inventory)
3. Sensory (visual-auditory-kinesthetic) models (several inventories)
4. Social interaction models (Grasha-Reichmann Student Learning Style Scales and Learning Preference Scales)
5. Howard Gardner’s multiple intelligences model (several inventories)
6. John Biggs’ approaches to learning model (Study Process Questionnaire)

These models are not mutually exclusive or necessarily complementary (Vincent & Ross, 2001). The field dependence/field independence model, for example, is similar to the Jungian

sensing-intuition scale, and the social interaction models are similar to the Jungian introversion-extroversion scale.

Some instruments (Index of Learning Styles, Productivity Environmental Preference Survey, and some of the instruments listed above) draw on multiple models.

### Is the concept of learning style valid and useful?

Vincent and Ross (2001) note that most professional educators “agree that learning styles exist and acknowledge the significant effect that learning styles have on the learning process.” The concept of learning styles makes sense intuitively. It is apparent, for example, that some people prefer reading books rather than listening to them on tape and vice versa, and that some people prefer working alone rather than working with others and vice versa (Curry, 1987). Indeed, some learning preferences (e.g., a preference for a quiet background) seem so self-evident that it may not be necessary to have a validated instrument to assess those preferences. As Nagy (1995) notes, “Little can be said about preference questions that ask, for example, what time of day a student prefers to study, except to wonder if such information requires the expense of a standardized test.”

Learning style advocates point to a number of validating studies. Swanson (1995), for example, cites numerous studies identifying cultural differences in learning styles, and the discussions of individual instruments that follow include other examples of validating studies.

Critics, however, point out that for a learning style theory to be valid and useful, it must be shown that students learn more effectively when their learning styles are accommodated, and only a limited number of studies have shown this. Some therefore feel that the usefulness or validity of learning style models and instruments has not been definitively

established (Bonham, 1988a; Bonham, 1988b; Kavale & Forness, 1987; Rayner & Riding, 1997).

While unvalidated instruments should not be used to make potentially harmful decisions about students, pedagogy, curriculum, etc., they may be used to help students gain self-awareness, provided that students have the opportunity to complete several instruments, so they do not take the results of any one instrument too seriously.

A particular concern (Grasha, 1990; Stellwagen, 2001) is that most learning style theories label or pigeonhole students into a few discrete, quantitative, often dichotomous categories, rather than recognizing that individuals develop and practice a qualitative mixture of learning styles that evolve as they learn and grow and that vary by discipline (Silver, Strong, & Perini, 1997). In some instances, pigeonholing begets a risk of stereotyping cultural groups.

### **How can faculty members use information on learning styles to help students learn?**

Faculty members first should understand that students use a variety of approaches to learning that may not match their own. Schroeder (1993) reports that over 75% of faculty members prefer the Myers-Briggs intuitive style and learn best through abstract concepts, ideas, and theories, compared to just 40% of entering students and 25% of the general population. Most students, in contrast, prefer the sensing style and learn best through concrete, practical, structured, and sequential experiences.

Anderson and Adams (1992) and Wilson (1998) urge faculty members to use a flexible variety of approaches to help students learn; Montgomery and Groat (2002), and Vincent and Ross (2001) offer specific suggestions. In addition to offering the usual lectures and readings, faculty can engage their students' senses and give them an idea of structure by providing visual aids such as bulleted lists, charts, and diagrams; written outlines or study guides of key points; structured opportunities for group interaction; practical "real world" examples; and a variety of assignment formats. They can also

try to get to know their students, help their students get to know them, and provide plenty of feedback. Most good teachers, of course, already do this instinctively.

Claxton and Murrell (1987) further recommend that faculty members:

1. Participate in workshops and other professional development opportunities that help them better understand the importance of learning style and its role in improving students' learning.
2. Engage in classroom research that investigates how information on learning styles can improve their teaching.
3. Create curricular experiences that help students learn how to learn by raising their awareness of their own preferences and strengths and developing strategies for succeeding in courses taught by faculty members whose styles differ from their own.

### **Should faculty members aim to accommodate each student's learning style?**

Faculty members should not try to accommodate individual learning styles for several reasons.

1. Models of teaching and learning styles are not yet sufficiently validated to be able to determine definitively how each student learns best and customize instruction accordingly (Zarghani, 1988).
2. As Gregorc (cited in Wilson, 1998) notes, "attempting to teach to all students' styles can quickly cause a teacher to burn out." Gardner (1996) states that "there is no point in assuming that every topic can be effectively approached in at least seven ways, and it is a waste of effort and time to attempt to do this."
3. While students should use their "strong" learning styles to best advantage, it's just as appropriate for them to develop their abilities to use other learning styles (Grasha, 1990; Montgomery & Groat, 2002) and to work with faculty whose styles differ from their own.

## How might students learn about their own learning styles?

Because any one instrument is an incomplete, imperfect assessment of learning style, students should not let any one instrument dictate their learning styles (Bonham, 1988a). Instead, they should be encouraged to develop their own sense of their learning styles, using multiple learning style inventories as clues rather than as definitive determinations. Crowe (2000) and Grasha (1990) suggest that students could be asked to:

- ❑ Write a paragraph or two explaining how they learn best.
- ❑ Complete at least two learning style inventories (perhaps chosen from among the Keirsey Temperament Sorter, the VARK Questionnaire, Owens and Barnes' Learning Preference Scales, the Multiple Intelligences Developmental Assessment Scales, and the Study Process Questionnaire) and compare the results with their self-description.
- ❑ After completing a course project, write a reflection on what and how they learned from the project.
- ❑ Use all of the above to develop a refined statement of how they learn best, along with a list of study/learning strategies they could use to take best advantage of their own particular learning style and to help them learn in situations where they must use approaches that do not correspond with their style.

## References

- Anderson, J. A. (2001). Developing a learning/teaching style assessment model for diverse populations. In Linda Suskie (Ed.), *Assessment to promote deep learning: Insight from AAHE's 2000 and 1999 Assessment Conferences*. Washington: American Association for Higher Education.
- Anderson, J. A., & Adams, M. (1992). Acknowledging the learning styles of diverse student populations: Implications for instructional design. In L. L. B. Border & N. V. Chism (Eds.), *Teaching for diversity*. San Francisco: Jossey-Bass.
- Bonham, L. A. (1988a). Learning style use: In need of perspective. *Lifelong Learning: An Omnibus of Practice and Research*, 11(5), 14-17, 19.
- Bonham, L. A. (1988b). Learning style instruments: *Let the buyer beware*. *Lifelong Learning: An Omnibus of Practice and Research*, 11 (6), 12-16.
- Claxton, D. S., & Murrell, P. (1987). *Learning styles: Implications for improving educational practices* (Report No. 4). Washington: Association for the Study of Higher Education.
- Cornwell, J. M., & Manfredro, P. A. (1994, Summer). Kolb's learning style theory revisited. *Educational & Psychological Measurement*, 54 (2), 317-328.
- Crowe, R. (2000). *Know your student's learning style: The missing link in the lecture v. active learning issue*. Paper presented at the NISOD Conference, Austin, TX.
- Curry, L. (1987). *Integrating concepts of cognitive or learning style: A review with attention to psychometric standards*. Ottawa: Canadian College of Health Service Executives.
- Dunn, R., & Griggs, S. (1995). A meta-analytic validation of the Dunn and Dunn model of learning-style preferences. *Journal of Educational Research*, 88 (6), 353-362.
- Gardner, H. (1996, Apr.). Multiple intelligences: Myths and messages. *International Schools Journal*, 15 (2), 8-22.
- Grasha, T. (1990). The naturalistic approach to learning styles. *College Teaching*, 38 (3), 106-114.
- Hayes, J. & Allinson, C. W. (1997, May-June). Learning styles and training and development in work. *Educational Psychology*, 17 (1/2), 185-194.
- Kavale, K. & Forness, S. (1987). Style over substance: Assessing the effectiveness of modality testing and teaching. *Exceptional Children*, 54, 228-239.
- Kavale, K. A., Hirshoren, A., & Forness, S. R. (1998, Spring). Meta-analytic validation of the Dunn and Dunn model of learning-style preferences: A critique of what was Dunn. *Learning Disabilities Research and Practice*, 13 (2), 75-80.
- Keefe, J. W. (Ed.) (1979). *Student learning styles: Diagnosing and prescribing programs*. Reston, VA: National Association of Secondary School Principals.
- Klein, P. D. (1997, Fall). Multiplying the problems of intelligence by eight: A critique of Gardner's theory. *Canadian Journal of Education*, 22 (4), 377-394.

- Lewthwaite, B., & Dunham, P. H. (1999). *Enriching teaching scholarship through learning styles*. Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education, Washington.
- Montgomery, S. M., & Groat, L. N. (2002). *Student learning styles and their implications for teaching*. Ann Arbor: Center for Research on Learning and Teaching, University of Michigan.
- Nagy, P. (1995). Review of the Learning Style Profile. In J. C. Conoley & J. C. Impara (Eds.), *The Twelfth Mental Measurements Yearbook*. Lincoln, NE: Buros Institute of Mental Measurements.
- O'Neil, J. (1990). Making sense of style. *Educational Leadership*, 48 (2), 4-9.
- Presland, J. (1994). Learning styles and continuous professional development. *Educational Psychology in Practice*, 10, 179-184.
- Rayner, S. & Riding, R. (1997). Towards a categorization of cognitive styles and learning styles. *Educational Psychology*, 17 (1/2), 5-29.
- Ruble, T. L., & Stout, D. E. *A critical assessment of Kolb's learning style inventory*. ERIC Document Reproduction Service No. ED 377221.
- Rule, D. L., & Grippin, P. C. (1988). *A critical comparison of learning style instruments frequently used with adult learners*. Paper presented at the Annual Conference of the Eastern Educational Research Association, Miami Beach.
- Schroeder, C. C. (1993, Sept./Oct.) New students – New learning styles. *Change*.
- Seidel, L. E., & England, E. M. (1997). *Gregorc's cognitive styles: Preference for instructional and assessment techniques in college students*. Poster presented at the Annual Convention of the American Psychological Society, Washington.
- Sewall, T. J. (1986). *The measurement of learning style: A critique of four assessment tools*. University of Wisconsin at Green Bay, Assessment Center. ERIC Document Reproduction Service No. ED 267247.
- Shearer, B. (n.d.). *Assessment of the multiple intelligences: Theoretical issues*. Multiple Intelligences Special Interest Group of the American Educational Research Association. Available online: <http://www.geocities.com/Athens/Column/7568/assessment.htm>.
- Silver, H., Strong, R., & Perini, M. (1997, Sept.) Integrating learning styles and multiple intelligences. *Educational Leadership*, 55 (1), 22-27.
- Stellwagen, J. B. (2001, May/June). A challenge to the learning style advocates. *Clearing House*, 74 (5), 265-269.
- Swanson, L. J. (1995, July). *Learning styles: A review of the literature*. The Claremont Graduate School. ERIC Document Reproduction Service No. ED 387067.
- Thompson, P., et al. (1979, April). *Interrelationships among five cognitive style tests, student characteristics, and achievement*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco. ERIC Document Reproduction Service No. ED 174678.
- Traub, J. (1998, October 26). Multiple intelligence disorder. *The New Republic*.
- Vincent, A., & Ross, D. (2001, Summer). Learning style awareness: A basis for developing teaching and learning strategies. *Journal of Research on Technology in Education*, 33 (5).
- Wilson, V. A. (1998, Jan. 1). *Learning how they learn: A review of the literature on learning styles*. Muskingum College. ERIC Document Reproduction Service No. ED 427017.
- Zarghani, G. H. Z. (1988). *Identification of learning style strategies which enable college students with differing personality temperament to cope with learning blocks*. Unpublished doctoral dissertation, University of Nebraska, Lincoln.

© 2002, Copyright by Linda Suskie. Reproduced with permission.

# References

- Anastasi, A. & Urbina, S. (1996). *Psychological Testing, Seventh Edition*. Upper Saddle River, NJ :Prentice-Hall.
- Angelo, T. A. & Cross, P. K. (1993). *Classroom assessment techniques: A handbook for college teachers* (2nd ed.). San Francisco: Jossey-Bass.
- Angelo, T. A., Ewell, P.T., & Lopez, C. (2001). *Assessment at the millennium: Now what?* In L. Suskie (Ed.), *Assessment to Promote Deep Learning* (pp. 59-63). Washington, DC: American Association for Higher Education.
- Association of College & Research Libraries. (2000). *Information literacy competency standards for higher education*. Chicago: Author.
- Astin, A. (1991). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education*. Portland: Oryx Press and the American Council on Education.
- Banta, T. W., Lund, J. P., Black, K. E., & Oblander, F. W. (1996). *Assessment in practice: Putting principles to work on college campuses*. San Francisco: Jossey-Bass.
- Briggs-Myers, I. & McCaulley, M. (1992). *Manual and guide to the use of the Myers-Briggs Type Indicator*. Consulting Psychologists Press.
- Chickering, A. & Gamson, Z. F. (1991). Applying the seven principles for good practice in higher education. *New Directions for Teaching and Learning, No. 47*. San Francisco: Jossey-Bass.
- Gall, J. P., Borg, W. R. & Gall, M. D. (1998). *Applying educational research: A practical guide* (4th ed.). Boston: Allyn & Bacon.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Haladyna, T. M. (1999). *Developing and validation multiple choice test items*. Mahwah, NJ: Erlbaum.
- Hartman, V. F. (1995). Teaching and learning style preferences: Transitions through technology. *VCCA Journal, 9*, 2, 18-20.
- Huba, M. E. & Freed, J. E. (2000). *Learner-centered assessment on college campuses*. Needham Heights: Allyn & Bacon.
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Hillsdale, NJ: Prentice Hall.
- Krueger, R. A. & Casey, M. A. (2000). *Focus groups: A practical guide for applied research*. London: Sage.
- Maxwell, J. A. A. (1996). *Qualitative research design: An interactive approach*. London: Sage.
- Middle States Commission on Higher Education. (2006). *Characteristics of excellence in higher education: Eligibility requirements and standards for accreditation*. Philadelphia: Author.
- \_\_\_\_\_. (2003). *Developing research and communication skills: Guidelines for information literacy in the curriculum*. Philadelphia: Author.
- Morgan, D. L. (1997). *Focus Groups As Qualitative Research, Second Edition*. London: Sage.
- Palomba, C. A. & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, improving*. San Francisco: Jossey-Bass.
- Pascarella, E. T. & Terenzini, P. T. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco: Jossey-Bass.
- Schein, E. (1996). *Organizational culture and leadership*. San Francisco: Jossey-Bass.
- Silverman, D. J. (2001). *Interpreting qualitative data: Methods for analyzing talk, text and interaction*. (2nd ed.). London: Sage



- Sternberg, R. J. (1988). *The triarchic mind: A new theory of human intelligence*. New York: Viking Press.
- Stufflebeam, D. L. (2001). Evaluation models. *New Directions for Evaluation, No. 89*. San Francisco: Jossey-Bass.
- Suskie, L. (1996). *Questionnaire survey research: What works?* (2nd ed.). Tallahassee, Florida: Association for Institutional Research.
- Walvoord, B. E. & Anderson, V. J. (1998). *A tool for learning and assessment*. San Francisco: Jossey-Bass.