

Methods of Measuring Learning Outcomes and Value Added

	Education Researchers/IR	Faculty
Indirect Measures	<ul style="list-style-type: none">• Freshman/senior surveys• Alumni surveys• Graduation rates• Number of students progressing to advanced degrees	<ul style="list-style-type: none">• Grades¹• Course evaluations (during the semester and end-of-semester)• Concept questions, “muddy cards,”² and other in-class techniques
	<ul style="list-style-type: none">• Surveys of student attitudes about new pedagogy, curriculum, etc.; surveys asking students for reflections on their learning• Exit interviews	<p>NOTE: Methods spanning the vertical axis may be used by researchers or faculty.</p>
Direct Measures	<ul style="list-style-type: none">• Standardized tests of general education skills (e.g., Collegiate Learning Assessment)• Think-aloud protocols³	<ul style="list-style-type: none">• Grades¹• Standardized tests of disciplinary knowledge (e.g., Force Concept Inventory)
	<ul style="list-style-type: none">• Pre-post tests• Analysis of assignments designed to test conceptual understanding (e.g., concept maps, pro/con grids)⁴• Observations of students performing a task• Analysis of student work products (e.g., exams, essays, oral presentations)• Senior thesis• Portfolios compiled over course of undergraduate study	

Methods of Measuring Learning Outcomes and Value Added

The Methods of Measuring Learning Outcomes and Value Added Grid provides a way to categorize the range of methodologies that can be used to answer the broad question, what knowledge and abilities have students acquired from both their academic work and their co-curricular activities during their years in college? Taken together, the data collected using these methodologies can help assess the value added by a college education.

There are, in fact, a number of ways of classifying these methods (e.g., whether they yield quantitative or qualitative data; whether the approach will be experimental, quasi-experimental, or correlational). The Methods of Measuring Learning Outcomes and Value Added Grid, however, was built specifically to respond to concerns that have surfaced about accountability in the wake of the Commission on the Future of Higher Education. Much of the debate prompted by the Commission has been about what are—and what are not—credible means of measuring learning. Because the gold standard in assessment is to triangulate the data (i.e., use a mix of measures), the grid includes a range of methodologies. However, the methods do vary in the kind and degree of resources needed to employ them effectively. Therefore, the grid categorizes assessment methods along two continua:

- Whether the methodology collects data that are indirect or direct measures of student learning, and
- Whether the methodology is more easily designed and implemented, and the resulting data analyzed, by educational/institutional researchers, faculty, or both.

These two variables—whether measures are direct or indirect and who collects and analyzes the data—were chosen because they have strong political and resource implications that need to be taken into account in developing an assessment plan. A successful assessment policy is derived from choosing the best combination of methodologies that can be implemented using the available resources. The grid is designed to help institutions make decisions along those lines.

Direct and Indirect Measures

Educational research is tricky business. Methodologies that are used to measure student learning each have their own limitations and biases, and no method can be counted on to be completely error free. That is why best practice in educational research dictates triangulating the data. If several different sources of data are used, it increases the probability that the findings present an accurate picture. In other words, the strongest assessment programs will rely on a mix of direct and indirect measures.

Indirect measures include data from surveys of seniors and alumni, retention rates, graduation rates, number of students progressing to advanced degrees, etc. They allow

administrators, faculty, researchers, and consumers to infer the benefits to students from their years in college, but they cannot report with precision exactly what students have learned or what they are capable of doing as a result of their university education. Historically, these kinds of data have been collected by offices of institutional research, alumni offices, etc.

Direct measures provide more evidence of the increase in students' knowledge and abilities over a period of time. Standardized tests as, for example, the Collegiate Learning Assessment (CLA) are one kind of direct measure. While the CLA assesses general education skills, other standardized tests can measure specific disciplinary knowledge. The Force Concept Inventory, for instance, is used to determine students' understanding of concepts in mechanics.

Other examples of direct measures include assignments that ask students to perform some kind of conceptual task (e.g., create a concept map) or portfolios compiled over a course of study. It is important to emphasize that these student work products need to be systematically reviewed for evidence of learning in order for them to be of most use. For example, rubrics can be developed and used by groups of faculty or educational researchers to analyze papers, thesis, or portfolios in order to assess learning. Grades, of course, can also be a measure of learning although how the grades are determined and reported can sometimes undermine their usefulness.

Who Does the Assessment?

There are a number of ways that individual institutions have delegated responsibility for assessing learning outcomes and value added. Surveys, either locally developed ones or standardized instruments (e.g., the National Survey of Student Engagement), are often administered and analyzed by offices of institutional research. Similarly, some units at the university (e.g., the disability office) may run their own surveys. Institutions may have assessment and evaluation centers, put together committees on assessment and evaluation, locate A&E within a teaching and learning center, or call upon the expertise of faculty in the school of education. In these cases, assessment may be done by staff members, outside consultants, or both working independently or collaboratively.

Over the last two decades, there has also been increased interest in what is sometimes called "course-embedded assessment." This approach comes from the realization that faculty are assessing student learning all the time—it is what they do as part of their role as instructors. However, they may not be consciously looking at student work products for evidence of what their students have (or have not) learned. Patricia Hutchings and Lee Shulman write about the need for faculty to "go meta": that is, to analyze assignments, exams, etc. specifically for evidence of learning. This may require using more rigorous methods of investigation along with the usual ways of grading, which may, in turn, require training faculty or the expertise of an educational researcher.

Considerations in Using the Grid

As with many institutional efforts, choosing methods to assess learning outcomes and value added is a matter of balancing best practices against the constraints imposed by resources and political will. Methods that more directly measure student learning and yield the most rigorous results are usually the most time intensive and may require the expertise of educational researchers or outside consultants. Thus they can be relatively expensive. Faculty can use their classroom and grading practices to collect data that will contribute to assessing student learning, but that will require extra time and effort on their part. In addition, mechanisms to help them report their findings need to be put in place.

While there is no way around the fact that implementing and sustaining an assessment program will require resources, there are cost-efficient ways to put such a program in place. For example, qualitative methodologies like interviews or focus groups that tend to be resource intensive can be done on a representative sample. Graduate students from education or other social sciences can participate in research, perhaps for credit. And, as explained above, instruments like alumni surveys can do double duty.

It is hoped the Methods of Measuring Learning Outcomes and Value Added Grid can provide a snapshot into the issues that have to be taken into consideration when exploring assessment policies and programs.

Notes

¹ Grades, of course, do provide a measure of how much students have learned. However, the validity of grades as an assessment measure is dependent upon how systematically and rigorously assignments, exams, etc. are analyzed for evidence of student learning.

² Students are given several minutes at the end of class to write on 3"x5" index cards what the "muddiest point in the lecture" was for them.

³ As the name implies, think aloud protocols ask students to verbalize what they are thinking as they solve a problem. There are accepted methodologies for rigorously analyzing the data from think aloud protocols.

⁴ A distinction needs to be made between the assessment activity itself (e.g., concept maps, pro/con grid) and the method of measuring the outcome from that activity. For example, there are formal methods of analyzing concept maps that are probably not feasible for faculty because of the time and expertise involved in re-representing the map and running the analysis. Less time-consuming methods of analysis are possible, but, again, the data that results will not be as rigorous. (Anne Faye, personal e-mail, 5/16/07)