



Peer Review, Winter/Spring 2002

AAC&U *Peer Review*, Winter/Spring 2002, Volume 4, Number 2/3. *Peer Review* is a publication of the Association of American Colleges and Universities <http://www.aacu.org/index.cfm>. Copyright © 2002.

Reality Check

Beyond Confusion: An Assessment Glossary

By Andrea Leskes, vice president for education and quality initiatives, AAC&U

The articles and commentaries in this edition of *Peer Review* raise important issues for all college educators. Part of our professional responsibilities involves knowing if, what, and how well students learn what is being taught. Ideally, assessment would be a regular, inherent, and transparent part of all teaching and learning.

College professors regularly employ with comfort some types of assessment; they rightly point this out when challenged, explaining how they always evaluate student learning, using tests or homework assignments to do so. Assessment of this sort normally occurs within the confines of individual courses and can provide important information to both the student and the professor. However, higher education has less of a history of examining accomplishments that build cumulatively, over time, throughout a student's entire undergraduate career. Yet we acknowledge that many of the goals of college education are exactly these accomplishments (e.g., effective communication, ethical judgement, analytical acuity). These, too, are the complex accomplishments that the previous articles address.

Higher education lacks a common vocabulary about assessment; and individuals use terms in mutating ways to refer to varying levels of analysis. Some interpretations imply external oversight or control, an unpleasant idea for most college faculty. Miscommunication and mistrust result from this confused language and are likely to interfere with developing the kind of useful value added assessment proposed by Benjamin, Hersh, and Klein.

To shine a light, as Marc Chun so eloquently puts it, in the darkness and "where we should be looking," this "reality check" attempts to provide such a common vocabulary for the concepts used or implied in the preceding articles. Herewith, a glossary of educational assessment terms, within the college context and focusing on student learning.

Value added: the increase in learning that occurs during a course, program, or undergraduate education. Can either focus on the individual student (how much better a student can write, for example, at the end than at the beginning) or on a cohort of students (whether senior papers demonstrate more sophisticated writing skills-in the aggregate-than freshmen papers). Requires a baseline measurement for comparison.

Standards: sets a level of accomplishment all students are expected to meet or exceed. Standards do not necessarily imply high quality learning; sometimes the level is a lowest common denominator. Nor do they imply complete standardization in a program; a common minimum level could be achieved by multiple pathways and demonstrated in various ways. Examples: carrying on a conversation about daily activities in a foreign language using correct grammar and comprehensible pronunciation; achieving a certain score on a standardized test.

Formative assessment: the gathering of information about student learning-during the progression of a course or program and usually repeatedly-to improve the learning of those students. Example: reading the first lab reports of a class to assess whether some or all students in the group need a lesson on how to make them succinct and informative.

Summative assessment: the gathering of information at the conclusion of a course, program, or undergraduate career to improve learning or to meet accountability demands. When used for improvement, impacts the next cohort of students taking the course or program. Examples: examining student final exams in a course to see if certain

specific areas of the curriculum were understood less well than others; analyzing senior projects for the ability to integrate across disciplines.

Assessment for accountability: assessment of some unit (could be a department, program or entire institution) to satisfy stakeholders external to the unit itself. Results are often compared across units. Always summative. Example: to retain state approval, the achievement of a 90 percent pass rate or better on teacher certification tests by graduates of a school of education.

Assessment for improvement: assessment that feeds directly, and often immediately, back into revising the course, program or institution to improve student learning results. Can be formative or summative (see "formative assessment" for an example).

Qualitative assessment: collects data that does not lend itself to quantitative methods but rather to interpretive criteria (see the first example under "standards").

Quantitative assessment: collects data that can be analyzed using quantitative methods (see "assessment for accountability" for an example).

Direct assessment of learning: gathers evidence, based on student performance, which demonstrates the learning itself. Can be value added, related to standards, qualitative or quantitative, embedded or not, using local or external criteria. Examples: most classroom testing for grades is direct assessment (in this instance within the confines of a course), as is the evaluation of a research paper in terms of the discriminating use of sources. The latter example could assess learning accomplished within a single course or, if part of a senior requirement, could also assess cumulative learning.

Indirect assessment of learning: gathers reflection about the learning or secondary evidence of its existence. Example: a student survey about whether a course or program helped develop a greater sensitivity to issues of diversity.

Assessment of individuals: uses the individual student, and his/her learning, as the level of analysis. Can be quantitative or qualitative, formative or summative, standards-based or value added, and used for improvement. Would need to be aggregated if used for accountability purposes. Examples: improvement in student knowledge of a subject during a single course; improved ability of a student to build cogent arguments over the course of an undergraduate career.

Assessment of programs: uses the department or program as the level of analysis. Can be quantitative or qualitative, formative or summative, standards-based or value added, and used for improvement or for accountability. Ideally program goals and objectives would serve as a basis for the assessment. Example: how sophisticated a close reading of texts senior English majors can accomplish (if used to determine value added, would be compared to the ability of newly declared majors).

Assessment of institutions: uses the institution as the level of analysis. Can be quantitative or qualitative, formative or summative, standards-based or value added, and used for improvement or for accountability. Ideally institution-wide goals and objectives would serve as a basis for the assessment. Example: how well students across the institution can work in multi-cultural teams as sophomores and seniors.

Embedded assessment: a means of gathering information about student learning that is built into and a natural part of the teaching-learning process. Often uses for assessment purposes classroom assignments that are evaluated to assign students a grade. Can assess individual student performance or aggregate the information to provide information about the course or program; can be formative or summative, quantitative or qualitative. Example: as part of a course, expecting each senior to complete a research paper that is graded for content and style, but is also assessed for advanced ability to locate and evaluate Web-based information (as part of a college-wide outcome to demonstrate information literacy).

Local assessment: means and methods that are developed by an institution's faculty based on their teaching approaches, students, and learning goals. Can fall into any of the definitions here except "external assessment," for which is it an antonym. Example: one college's use of nursing students' writing about the "universal precautions" at multiple points in their undergraduate program as an assessment of the development of writing competence.

External assessment: use of criteria (rubric) or an instrument developed by an individual or organization external to the one being assessed. Usually summative, quantitative, and often high-stakes (see below). Example: GRE exams.

"High stakes" use of assessment: the decision to use the results of assessment to set a hurdle that needs to be cleared for completing a program of study, receiving certification, or moving to the next level. Most often the assessment so used is externally developed, based on set standards, carried out in a secure testing situation, and administered at a single point in time. Examples: at the secondary school level, statewide exams required for graduation; in postgraduate education, the bar exam.