

Why aren't grades enough?

(1) A discussion from North Carolina State University, retrieved from http://www.ncsu.edu/provost/academic_programs/uapr/FAQ/UAPRFAQwhatdifassessstudentvsprograms.html

What is the difference between assessing a program and assessing a student?

Some faculty members have wondered if they could just use student grades as data for assessing academic programs. They reason that if a program outcome is related to a particular course or assignment, the grades for that course or assignment should indicate the degree to which students are able to meet the outcome. That may be true in certain rare cases, but on the whole, student grades don't provide the best data for program assessment because the two kinds of assessment are different in important ways.

First, they have different purposes. In the classroom, we design and evaluate student assignments with the goal of helping students achieve the learning objectives specific to the class and the assignment. The purpose is to use assessment as a teaching tool, for guiding and testing student learning. Program assessment's purpose is to provide faculty with the information they need to improve their programs, to determine the degree to which the program is enabling students to meet program outcomes and to propose changes in the program as indicated.

Another difference is that student learning outcomes in individual classrooms are likely to be different from program outcomes. Learning outcomes are specific to a class and to the needs of students in that class and may change from teacher to teacher and semester to semester. They provide a framework for student learning. Program outcomes tend to be broader and more general, focused on what a program's courses have in common rather than the individual outcomes of each course. And because program assessment takes a broader view, there could be program outcomes that do not appear as specific learning outcomes in any class. This difference between outcomes means that the criteria by which a teacher assesses students and by which faculty assess a program are also likely to be different.

In addition to these differences, there is also the problem that grading processes vary across faculty members, across a single faculty member's courses, across semesters, and even across particular assignments in a single course. This natural variability makes it difficult to use grades as a reliable indicator of student abilities and thus as data for assessing a program. However, if a program gives all its majors a highly reliable and valid test or other assignment directly related to a program outcome, grades on that assignment could be used for program assessment.

The fact that student grades are typically inadequate as data for program assessment doesn't mean that student work itself--research projects, essay exams, lab reports, literature reviews--is also inadequate as data. Indeed, student work may be the very best data for many program outcomes. The difference is that faculty who are doing program assessment are most likely going to take a perspective on that work that differs substantially from the teacher of the course, applying different criteria for different purposes. Thus the same data may be used for both student and program assessment, but the way those data are used is not the same.

As faculty, we assess our students and we assess our programs. These are quite different activities with quite different purposes. However, they do have one critical goal in common: student learning. Student assessment is a teaching tool designed to encourage and evaluate student learning. Program assessment is an institutional tool designed to enhance academic programs in order to improve student learning. They each play an important role in creating a better learning environment for our students.

- Mike Carter
Associate Professor of English

(2) A discussion from California State University-San Bernadino, retrieved from <http://academic-affairs.csusb.edu/progs/assessment/forumf98.htm>

Why Aren't Course Grades Enough?

The Outcomes Assessment committee still is continually asked why course grades no longer are considered sufficient evaluation of student progress. "Don't grades have any meaning? " Well, of course, classroom grades are useful assessment tools, but the whole thrust of the assessment movement is to provide additional, more comprehensive, and more long-term evidence of student achievement. Why?

First, course grades represent evaluation of limited objectives that often are not related to program objectives. At best, grades are a type of "formative" assessment. There is too much inconsistency from section-to-section and term-to-term for grades to be a totally valid overall program assessment.

Secondly, there is an inherent conflict of interest when the instructor is the only evaluator of whether the students have met program objectives.

Third, we really need to know more about the long-term learning and accomplishments of our graduates. What do they know and what can they do at the end of the program? Finding out about their perceptions, attitudes and skills five years after graduation, on the job, in graduate school, or throughout the student's life is more difficult, but an area we are increasingly asked to provide information about.

The university assessment committee encourages departments to use classroom assessment measures and to collect data over time from key courses. We applaud the inclusion of this information on "formative" assessment in their overall assessment of the progress of their majors in meeting the stated goals and objectives of their program. Indeed, some programs have attempted to build their assessment plan around having common outcome expectations for core courses. This is a very intrusive, often difficult, and time-consuming approach, however. The Assessment Committee will not accept assessment plans which rely solely on the course instructor to set the criteria, collect and evaluate the data. Moreover, we need and expect some form of "summative" assessment to be done at the end of the program as well. The best assessment plans have multiple points and methods of assessment of student progress and accomplishment.

Retrieved from

<http://www.abet.org/Linked%20Documents-UPDATE/Assessment/Assessment%20Tips4.pdf>

Do Grades Make the Grade for Program Assessment?

Assessment Tips With Gloria Rogers

One of the most common questions from faculty when discussing outcomes assessment is, “We are already assessing students in courses; why can't we just use student grades as an indication of what our students know or can do?” *Grades represent the extent to which a student has successfully met the faculty member's requirements and expectations for a course.* Because many factors contribute to an assigned grade, it is almost impossible to make inferences about what a student knows or can do by only looking at the grades for a course.

In outcomes assessment at the program level, the primary question that needs to be answered is, “Can students demonstrate the ability to perform at an acceptable level in each of the program outcomes?” Program assessment focuses on providing evidence that students can demonstrate knowledge or skill directly linked to specific program outcomes. Grades *per se* **do not** provide that information.

One reason why course grades are not appropriate for program assessment is that *course content for any given subject may vary among faculty members teaching the same course.* When developing a course, the faculty member has to make many decisions. These include decisions about course content and course management. When deciding what topics and concepts to include in the course, the faculty member needs a clear view of how the course is aligned with other courses in the curriculum (e.g., introductory, elective, required, lower/ upper division, major, or service course). Decisions about course content are constrained by several factors: the amount of time the faculty member has to deliver the course, the knowledge and skills that students bring to the course, and the expectations other faculty have for learning brought to follow-on courses. Content may also vary with the individual faculty member's beliefs about what is important (topics, concepts, and levels of cognition students must demonstrate for each concept), the textbook chosen, and the faculty member's expertise and interests. Decisions are also made about how the course is managed, for instance the mode of delivery, number and types of tests, attendance policy, and grade structure. All of these variables contribute to the grades students receive, further confounding the ability to interpret the relationship of the grade to specific student knowledge or abilities.

Another reason why grades do not provide adequate information for program assessment is that *the grading policy in any course is dependent on the individual faculty member.* This is generally true even when there are multiple sections of the same course with common exams. Some faculty choose to give (or take away) points or partial credit for things that are not related to student learning (for example, attendance, class participation, and course evaluation). Some faculty grade on a curve; others have a fixed standard. Letter grades or numeric scores reflect the student's relative standing within the class or among other tests – relative to a set scale or relative to other students. They do not, however, tell the person interpreting the assigned grade/score what the student knows or can do, nor do they provide information about what topics or concepts he or she did not understand or how his or her learning can be improved.

Assessing program learning outcomes for the curriculum differs from assessing classroom learning outcomes in several ways, most notably the following:

When developing a curriculum, faculty collectively consider the objectives¹ their students will need to achieve after graduation. Once the objectives are identified, faculty decide what students should know or be able to do by the time of graduation in order to meet them. After the program outcomes² are set, the curriculum is developed/modified to represent a well articulated and aligned set of major and general education courses. Students are introduced to key concepts in the lower division courses. Then these concepts are applied in courses throughout the rest of the curriculum, as students move from knowing and understanding a concept to developing an ability to apply that knowing and understanding in various ways, in multiple settings. This process illustrates the cumulative learning effect of specific concepts and skills taught through individual courses. The assessment of program outcomes should reflect student-achievement-specific outcomes as a culmination of several classes and activities throughout the curriculum.

Just as faculty cannot include in a course everything associated with the subject matter of that course, a program cannot include in its curriculum every concept or skill set that is in the realm of possibilities for that curriculum. As in course preparation, several decisions need to be made by program faculty when determining the program outcomes to be assessed and managing the assessment process. These include deciding what learning outcomes are central to achieving the objectives, how many and what performance criteria³ will be assessed for each outcome, where in the curriculum students are getting the opportunity to demonstrate the desired performance criteria associated with the outcome, how often the outcomes will be assessed, how the outcomes are going to be assessed, and how the data gathered can be used for program improvement. As in classroom assessment, these decisions are constrained by factors related to the context of the program. Some of these factors include the nature of the objectives, type of institution/program, available resources and time, and make up of students served.

For program assessment, a numeric score that is **directly** linked to students' performance on a specific performance criteria can be used as evidence of program learning outcomes. For example, for the outcome, "Students have an understanding of ethical responsibility," one of the performance criteria could be, "Students will demonstrate the ability to evaluate the ethical dimensions of a problem in their engineering discipline." Faculty could develop a rubric to score student performance. A rubric is a descriptive rating scale with several different observable levels of performance possible for each performance criteria being assessed. Each performance level is described and assigned a numeric score (for example, 1 = exemplary, 2 = good, 3 = adequate, 4 = marginal, and 5 = unacceptable). The number of points on the scale will depend on the level of cognition or skill that the outcome requires – but that is a discussion for a later time. Reporting the percent of students who score at each of the levels provides data that are linked directly to the anticipated outcome and focus the evaluation and strategies for improvement. It is a numerical score that provides a great deal of information about what students know or can do – but it is not a grade.

Grades will continue to be an important part of the higher education culture and should be understood for what they represent. However, for program assessment, where the purpose of the assessment is to provide information about student learning at the program level, grades in courses generally have little use. This is not to say that students cannot demonstrate program outcomes in a classroom setting. But, the measure used to assess those outcomes should be used consistently, should reflect specific student knowledge or skills, and should be directly linked to specific performance criteria. *It is important to remember that the focus is not a score or grade, but the student knowledge or skill that is represented by that score or grade.*

¹Objective here is defined as the expected accomplishments of graduates during the first few years after graduation.

²Outcome here is defined as what a student knows or can do by the time of graduation.

³Performance criteria here are defined as the specific, measurable statements identifying the specific knowledge, skills, attitudes and/or behavior students must demonstrate as indicators of achieving the outcome.