

Course-Embedded Assessment and Use of Scoring Rubrics

1. Course-Embedded Assessment. Assessment practices embedded in academic courses generate information about what and how students are learning within the program and classroom environment. Course-embedded assessment takes advantage of already existing curricular offerings by using standardized data instructors already collect or by introducing new assessment measures into courses. The embedded methods most commonly used involve the development and gathering of student data based on questions placed in course assignments. These questions, intended to assess student outcomes, are incorporated or embedded into final exams, research reports, and term papers in senior-level courses. The student responses are then evaluated by two or more faculty to determine whether or not the students are achieving the prescribed educational goals and objectives of the department. This assessment is a separate process from that used by the course instructor to grade the exam, report, or term paper.

There are a number of advantages to using course-embedded assessment. First, student information gathered from embedded assessment draw on accumulated educational experiences and familiarity with specific areas or disciplines. Second, embedded assessment often does not require additional time for data collection, since instruments used to produce student learning information can be derived from course assignments already planned as part of the requirements. Third, the presentation of feedback to faculty and students can occur very quickly creating a conducive environment for ongoing programmatic improvement. Finally, course-embedded assessment is part of the curricular structure and students have a tendency to respond seriously to this method. Departments at other research institutions using embedded assessment include general education programs, classics, economics, English, film studies, geography, fine arts, history, kinesiology, philosophy, political science, physics, and religious studies.

Source: <http://www.wisc.edu/provost/assess/manual/manual2.html#a2>

2. Scoring Rubrics. (Excerpts from Mary J. Allen, Ph.D. Director, California State University Institute for Teaching and Learning. March 22, 2004). Program assessment is an ongoing, formative process in which faculty specify program learning objectives, verify that the curriculum aligns with these objectives, collect assessment data, and use findings to improve student learning (Allen, 2004). Assessment data frequently are based on analyzing student products or behaviors, such as exam responses, projects, portfolios, or recitals. Scoring rubrics are versatile tools for simplifying this review by clearly specifying assessment criteria.

Rubrics are explicit schemes for classifying products or behaviors into categories that vary along a continuum. They can be used to classify virtually any product or behavior, such as essays, research reports, oral presentations, and group activities.

A wide variety of rubrics have been developed, and it is often easier to adapt an already-existing rubric than to create one from scratch. For links to online rubrics, go to <http://www.calstate.edu/acadaff/sloa/>. Many quality rubrics have been created for use in K-12 education, and they can be easily adapted for higher education. As you review available rubrics, don't look for one that you can copy and immediately use in your own work. Look for formats, language, and dimensions that might be useful components of a rubric that you tailor for your specific needs.

Rubrics have many strengths:

- Complex products or behaviors can be examined efficiently. Faculty have many demands on their time, and assessment activities should be structured to use that time effectively. Rubrics

focus raters on the learning objectives being assessed, allowing them to tune out extraneous variables. For example, if faculty are analyzing a set of lab reports to assess students' ability to statistically analyze data, the rubric should help them ignore other aspects of the reports, such as the quality of the lit review and written communication skills.

- Well-trained reviewers apply the same, agreed-upon standards to the products being reviewed. This generates data that are likely to be reliable and valid.
- Summaries of results reveal patterns of student strengths and areas of concern. These assessments allow us to identify learning objectives that require increased attention.
- Rubrics are criterion-referenced, rather than norm-referenced. Raters ask, "Did the student meet the criteria for level 5 of the rubric?" rather than "How well did this student do compared to other students?" This is important for program assessment because you want to learn how well students have met your standards.
- Faculty might feel overwhelmed when faced with assessment mandates, but they are not the only ones who can assess student work. Sometimes ratings can be done by students to assess their own work, or they can be done by others, e.g., peers, fieldwork supervisions, or visiting artists or scholars.

Rubrics can be used for grading *and* assessment. For example, points can be assigned and used for grading, as shown below, and the categories (below expectation, satisfactory, and exemplary) can be used for assessment. Faculty who share an assessment rubric might assign points in different ways, depending on the nature of their courses, and they might decide to add more rows for course-specific criteria or comments. In this way, assessment data can be collected while faculty are grading, faculty control how their own grades are assigned, and data are aggregated across relevant courses and faculty to assess the program. Walvoord and Anderson's provide many useful examples in *Effective Grading* (1998).

Analytic Rubric for Oral Presentations				
	Below Expectation	Satisfactory	Exemplary	Score
Organization	No apparent organization. Evidence is not used to support assertions. (0-2)	The presentation has a focus and provides some evidence which supports conclusions. (3-5)	The presentation is carefully organized and provides convincing evidence to support conclusions. (6-8)	
Content	The content is inaccurate or overly general. Listeners are unlikely to learn anything or may be misled. (0-2)	The content is generally accurate, but incomplete. Listeners may learn some isolated facts, but they are unlikely to gain new insights about the topic. (5-7)	The content is accurate and complete. Listeners are likely to gain new insights about the topic. (10-13)	

Style	The speaker appears anxious and uncomfortable, and reads notes, rather than speaks. Listeners are largely ignored. (0-2)	The speaker is generally relaxed and comfortable, but too often relies on notes. Listeners are sometimes ignored or misunderstood. (3-6)	The speaker is relaxed and comfortable, speaks without undue reliance on notes, and interacts effectively with listeners. (7-9)	
Total Score				

Steps that might be used to create a rubric and some additional suggestions for integrating rubrics into courses are provided at http://www.calstate.edu/acadaff/sloa/links/using_rubrics.shtml.

Another template: http://projects.edtech.sandi.net/memorial/franchise/business_plan_rubric.htm

One of the most positive aspects about program assessment is that it provides a framework for faculty discussions about teaching and learning, and, in my experience, group reviews of student products using rubrics can be an effective precursor of these discussions. Like any effective process, planning is required. The leadership team identifies appropriate student products that reflect student mastery of the learning objectives being assessed, collects these products, develops and pilot tests a rubric, and selects exemplar products differing in quality. The session generally begins with a review of the assessment effort, the products, and the rubric, then readers are “calibrated” by discussing and reaching consensus on ratings for the exemplar products.

Generally two reviewers apply the rubric to each product so that inter-rater agreement can be examined. Discrepancies often are resolved by using a third reader or by having paired reviewers reach consensus on each product. Once data are collected and summarized, the group can discuss what results mean (Have our students mastered our learning objectives at a satisfactory level?), who should be told the results, and implications for changes in curriculum, pedagogy, or student support.

Source: <http://cai.cc.ca.us/workshops/RubricsByMaryAllen.doc>

A STATEMENT REGARDING GRADING AND PROGRAM ASSESSMENT: For the purpose of conducting systematic assessment, grades awarded students are insufficient measures of student learning. They do not alone tell faculty precisely what students have learned. And, of course, standards for assigning grades can vary from faculty member to faculty member and department to department, undermining the idea that all grades of a given value are comparable. More importantly, grades are awarded for individual courses and not entire programs, and the purpose of assessment is to allow units to determine whether they are achieving their educational objectives.

Suggested Readings

Allen, M. J. (2004). *Assessing academic programs in higher education*. Bolton, MA: Anker.

Walvoord, B. E., & Anderson, V. J. (1998). *Effective grading*. San Francisco, CA: Jossey-Bass.

http://www.ncsu.edu/provost/academic_programs/uapr/FAQ/UAPRFAQwhatdifassessstudentvsprogrms.html