

Assessment



Galveston College *A Beacon of Light Guiding Lifelong Learning*

Two boys are walking down the street. The first boy says, “I’ve been really busy this summer. I’ve been teaching my dog to talk.”

His friend responds, “Wow! I can’t wait to have a conversation with your dog.”

The first boy shakes his head. “I said I’ve been teaching him. I didn’t say he learned anything.”

**from Mary J. Allen (2004).
Assessing Academic Programs in Higher Education
Anker Publishing, Inc.: Bolton, MA.**



TWO DEFINITIONS TO KNOW:

Assessment: *A systematic process of gathering and interpreting information to learn how well your unit is performing, and using that information to modify your operations in order to improve that performance.*

Outcome: With respect to a *learning outcome*, what do you want a student to *know* or *do* at the completion of an activity or course of study (knowledge, skills, or values)? With respect to an *administrative or other support unit outcome*, how will the *student's life or campus experience be improved* as a result of the services provided?



Assessment can be classified in different ways:

Formative: 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.

Or

Summative: The gathering of information at the conclusion of a course, program, or undergraduate career to improve learning or to meet accountability demands.

Qualitative: Collects data that does not lend itself to quantitative methods but rather to interpretive criteria

Or

Quantitative: Collects data that can be analyzed using quantitative methods



Norm-referenced: An assessment where student performance or performances are compared to a larger group. The larger group or "norm group" may be a national sample or a peer group of institutions. The SAT is a norm-referenced assessment. The purpose of a norm-referenced assessment is usually to sort students and not to measure achievement towards some criterion of performance.

Or

Criterion-Referenced: Assessment that measures student knowledge and understanding in relation to specific standards or performance objectives, not in relation to other students; all students may earn the highest rating if all meet the established performance criteria.



Direct: Gathers evidence about student learning based on student performance that demonstrates the learning itself. Can be value-added, related to standards, qualitative or quantitative, embedded or not, using local or external criteria. Examples are written assignments, classroom assignments, presentations, test results, projects, logs, portfolios, and direct observations.

Or

Indirect: Acquiring evidence about how students feel about learning and their learning environment, rather than actual demonstrations of outcome achievement. Examples include surveys, questionnaires, course evaluations, interviews, focus groups, and reflective essays.



3 Challenges for GC

1. Over-reliance on exam scores, in general, and multiple-choice questions, in particular.
2. Lack of emphasis on program-level and institutional-level outcomes
3. Relative absence of “Use of Results for Improvement”



What is the difference between assessment and grading?

Grading:

- may or *may not* reflect achievement of specific program or general education outcomes.
- is *course-specific*. That is, grades are assigned for courses, not programs, so they have little usefulness in assessing programs.
- is often *instructor-specific*, with different instructors assigning different grades for comparable work.
- usually includes other factors—like attendance, participation, & effort—that are not measures of learning outcomes.
- typically involves a comparative standard of measurement—how is the student doing in relation to other students—establishing a competitive relationship among those receiving grades (i.e., *normative-referenced*, not *criterion-referenced*)



Assessment:

- The goal is to *improve student learning* (“value added”). It tells us how WE are doing with respect to our students.
- It is *outcome-specific*, with the assessment methodology or measurement chosen to *directly measure* the extent to which a particular outcome or competency is achieved.
- Systematically examines patterns of student learning across courses and programs. [So, it can be used to validate *program-level &/or institutional-level* outcomes.]
- Measures student growth and progress on an individual basis (*criterion-referenced*). You may use sampling to assess students, rather than evaluating each student, but your goal is to *conclude* or *infer* the *extent to which* your students *as a whole* have mastered the outcomes [e.g, a cross-section or stratified sample of students may be tested to draw inferences on a larger *cohort* of students].



If You MUST Use Grades as an indicator of Student Learning:

- 1. Grades would first have to be decomposed into the components that are indicators of learning outcomes and those that are indicators of other behaviors.**
- 2. Grades would have to be based on *clearly articulated criteria* that are *consistently applied*. (i.e., across courses, programs, & faculty)**
- 3. Separate grades or sub-scores would have to be computed for the major components of knowledge and skills so that evidence of students' specific areas of strength and weakness could be identified. (as opposed to their overall course grade)**
- 4. As a rule, NEVER use course grade distributions to document or validate the acquisition of specific student learning outcomes!! This doesn't establish proof of ANYTHING, other than the instructor's approach to teaching & grading.**



OK, If I'm not supposed to use course grades, what CAN I use to measure student performance?

Pre-test/Post-test. This is where you administer a “quiz” at the first class session of the course to determine what your students know about the subject matter you are about to teach them, and the same (or similar test) at the end to determine what they have learned. Note: NOT the Final Exam, unless it specifically corresponds to the pre-test. (This is a value-added approach.)

Rubrics. A rubric is a specific set of criteria that clearly defines for both student and teacher what a range of acceptable and unacceptable performance looks like. Criteria define descriptors of ability at each level of performance and assign values to each level. A simple 3-point rubric might be: “Does not meet expectations,” “Meets expectations,” and “Exceeds expectations.”



Alternatives to course grades (Cont.):

Capstone courses. If your academic program includes a final course that sort of “pulls all the threads together,” this is an excellent opportunity to assess the students’ grasp of both program and general education learning outcomes. For example, you might assign a paper or project that brings together the program learning outcomes and general education competencies. Using rubrics for the program and general competencies will allow the instructor (or, better, a group of evaluators) to assess the extent to which the student has mastered the outcomes.

Authentic Assessment: This simply means to assess real world skills and knowledge in real world settings. That is, you are assessing the *application* of knowledge and skills acquired. The capstone course project is an excellent example of this. In workforce-related programs, you might use a rubric to assess knowledge and skills demonstrated in clinical or internship courses or settings.



Alternatives to course grades (Cont.):

Licensure exams. It is hard to beat a licensure examination for validating the program learning outcomes, if not the general education outcomes. It is external in origin, global in its perspective, and directly tied to the expectations of the profession or career specialization that has been the focus of the academic program.

English Proficiency Examinations. This might be an impromptu essay students are asked to write, which can be evaluated by a handful of faculty to determine if students have adequate writing skills.

Portfolios. A portfolio is a systematic and organized collection of a student's work that exhibits to others the direct evidence of a student's efforts, achievements, and progress over a period of time.



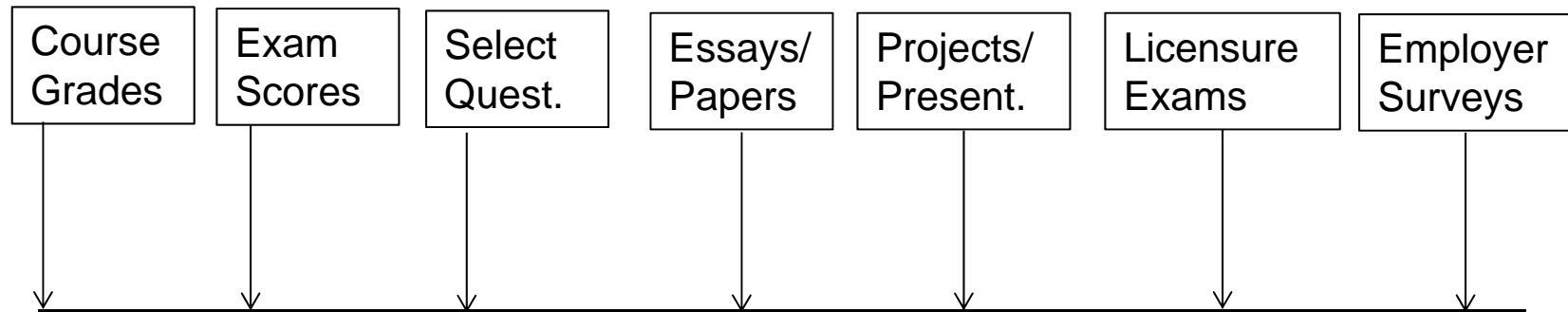
Alternatives to course grades (Cont.):

Norm-referenced exams. Examinations like the Measure of Academic Proficiency & Progress (a product of ETS) measure the extent to which students have mastered the general education competencies. A cross-section of students might be selected to take the test. The results would not affect student grades in any way, but rather provide a baseline of data to compare results over time. (This is also *criterion-referenced*, in that its comparison to students in a norm group is based on mastery of the competencies.)

Surveys. Graduate and Employer surveys are ways to determine the effectiveness of an academic or workforce program. Employers, for example, would be asked to assess the job readiness of the graduates. Graduates might be asked their level of satisfaction with their academic program and the extent to which it prepared them for the workforce and employer expectations.



Assessment as a Continuum...



Galveston College

A Beacon of Light Guiding Lifelong Learning

Examples of Learning Outcomes & Measurements

Outcome: Prepare & present an effective persuasive presentation appropriate to a specified audience, using the introduction, body & conclusion format, as well as an appropriate application of presentation technology.

Bad measurement: 70% of learners will achieve a grade of “C” or better in SPCH1311.

Good measurement: 75% of students will average at least a 3 on a 5-point Likert scale assessment on a final class presentation scored by a panel of evaluators.



Examples (Continued)

Outcome: Students will learn basic laboratory skills and safety rules.

Measurement: Students will satisfactorily complete the laboratory reports with a score of 70% or higher.

Outcome: Students will demonstrate an ability to identify, analyze, and discuss scarcity and supply and demand.

Measurement: Learners will score 70% or better on a computer-generated presentation to be critiqued on the basis of a faculty rubric.



Examples (Continued)

Outcome: The student will demonstrate basic competency with introductory geometry, constructions, congruence, and similarity.

Measurement: Separate Exam for each learning outcome with attainment of 70%.

Outcome: Understand fluid dynamics

Measurement: Student will answer at least 12 out of 20 questions (60%) on fluid dynamics.



Examples (Continued)

Outcome: Demonstrate an understanding of the mathematics of research methodology of the social behaviors of diverse world cultures.

Measurement: Learners will score 70% or better on a group research project that demonstrates an understanding of the mathematics of research methodology for social behaviors in a diverse world.

Outcome: Use technology to locate major works in the humanities.

Measurement: A total of 10 focused one-page written entries collected into a WebCT humanities portfolio, scoring 70% or better.



Examples (Continued)

Outcome: Students will produce compositions that make use of primary and secondary source materials that propose an understanding of short stories and novels.

Measurement: Student compositions will contain 1,000 words and evidence of at least 3 secondary source documents related to at least 1 primary source document. 70% of the secondary source material will respond to primary source material.

Outcome: Demonstrate an ability to critically analyze 1 of the 5 accepted methods of sociological research.

Measurement: Students will achieve a grade of 70% or better on a written term project portfolio scored by a faculty designated rubric identifying research methods.



So, what would you say are the characteristics of a good learning outcome?

And of a good measurement?

A good benchmark or criteria for success?

