

Assessment Terminology for Gallaudet University

A Glossary of Useful Assessment Terms

It is important that faculty and staff across all facets of campus agree on the meaning of terms that will be used in assessing courses and programs. The intent of this glossary is, therefore, to promote a clear and common definition of terms rather than promote the use of jargon. Nonetheless we solicit faculty and staff dialogue on these definitions and invite suggestions that aim to promote a common understanding about assessment.

Accountability

The obligation placed on an educational institute by public officials, employers, and taxpayers for school officials to prove that money invested in education has led to measurable learning.

Accountability is often viewed as an important factor in education reform. An assessment system connected to accountability can help identify needs so that resources can be equitably distributed.

The responsibility of an agency to its sponsors and clientele for accomplishing its mission with prudent use of its resources. In education, accountability is currently thought to require measurable proof that faculty and institutions are teaching students efficiently and well, usually in the form of student success rates on various tests.

Accreditation

Official recognition that an institution meets required standards. Gallaudet is accredited by a voluntary regional accreditation association, the Middle States Commission on Higher Education.

Achievement Test

A standardized test designed to efficiently measure the amount of knowledge and/or skill a person has acquired, usually as a result of classroom instruction. Such testing produces a statistical profile used as a measurement to evaluate student learning in comparison with a standard or norm.

Affective

Outcomes of education involving feelings more than understanding: likes, pleasures ideals, dislikes, annoyances, values.

Alternative Assessment

Alternatives to traditional, standardized, norm- or criterion-referenced traditional paper and pencil testing. An alternative assessment might require students to answer an open-ended question, work out a solution to a problem, demonstrate skill, or in some way produce work rather than select an answer from choices on a sheet of paper. Portfolios and instructor observation of students are also alternative forms of assessment. (Also Assessment Alternatives)

Analytic Scoring

A type of rubric scoring that separates the whole into categories of criteria that are examined one at a time. Student writing, for example, might be scored on the basis of grammar, organization, and clarity of ideas. Useful as a diagnostic tool. An analytic scale is useful when there are several dimensions on which the piece of work will be evaluated. (See **Rubric**.)

Aptitude Test

A test intended to measure the test-taker's innate ability to learn, given before receiving instruction.

Assessment

“Assessment is an ongoing process aimed at understanding and improving student learning. It involves making our expectations explicit and public; setting appropriate criteria and standards for learning quality; systematically gathering, analyzing, and interpreting evidence to determine how well performance matches those expectations and standards, and using the resulting information to document, explain and improve performance” (Tom Angelo, 1995)

“Assessment is the systematic collection, review and use of information about educational programs undertaken for the purpose of improving student learning and development.” (Palomba & Banta, 1999)

The systematic collection, review, and use of information about educational programs undertaken for the purpose of improving student learning and development.

The Latin root *assidere* means to sit beside. In an educational context, the process of observing learning; describing, collecting, recording, scoring, and

interpreting information about a student's learning. At its most useful, assessment is an episode in the learning process; part of reflection and autobiographical understanding of progress. Traditionally, student assessments are used to determine placement, promotion, graduation, or retention.

In the context of institutional accountability, assessments are undertaken to determine the effectiveness of academic programs, etc. In the context of school reform, assessment is an essential tool for evaluating the effectiveness of changes in the teaching-learning process.

Assessment Literacy

The possession of knowledge about the basic principles of sound assessment practice, including terminology, the development and use of assessment methodologies and techniques, familiarity with standards of quality in assessment. Increasingly, familiarity with alternatives to traditional measurements of learning.

Assessment Task

An illustrative task or performance opportunity that closely targets defined instructional aims, allowing students to demonstrate their progress and capabilities.

Authentic Assessment

Evaluating by asking for the behavior the learning is intended to produce. The concept of model, practice, feedback in which students know what excellent performance is and are guided to practice an entire concept rather than bits and pieces in preparation for eventual understanding. A variety of techniques can be employed in authentic assessment.

The goal of authentic assessment is to gather evidence that students can use knowledge effectively and be able to critique their own efforts. Tasks used in authentic assessment are meaningful and valuable, and are part of the learning process.

Authentic assessment can take place at any point in the learning process. Authentic assessment implies that tests are central experiences in the learning process, and that assessment takes place repeatedly. Patterns of success and failure are observed as learners use knowledge and skills in

slightly ambiguous situations that allow the assessor to observe the student applying knowledge and skills in new situations over time.

Benchmark

Student performance standards (the level(s) of student competence in a content area.)

An actual measurement of group performance against an established standard at defined points along the path toward the standard. Subsequent measurements of group performance use the benchmarks to measure progress toward achievement.

Examples of student achievement that illustrate points on a performance scale, used as exemplars. (See **Descriptor, Cohort, Criteria/Standards.**)

Bloom's Taxonomy of Cognitive Objectives

Benjamin Bloom originated this taxonomy for categorizing level of abstraction of questions that commonly occur in educational settings. The taxonomy provides a useful structure in which to categorize test questions, since professors will characteristically ask questions within particular levels, and if you can determine the levels of questions that will appear on your exams, you will be able to study using appropriate strategies. There are six levels arranged in order of increasing complexity (1=low, 6=high):

1. Knowledge: Recalling or remembering information without necessarily understanding it. Includes behaviors such as describing, listing, identifying, and labeling.
2. Comprehension: Understanding learned material and includes behaviors such as explaining, discussing, and interpreting.
3. Application: The ability to put ideas and concepts to work in solving problems. It includes behaviors such as demonstrating, showing, and making use of information.
4. Analysis: Breaking down information into its component parts to see interrelationships and ideas. Related behaviors include differentiating, comparing, and categorizing.
5. Synthesis: The ability to put parts together to form something original. It involves using creativity to compose or design something new.
6. Evaluation: Judging the value of evidence based on definite criteria.

Behaviors related to evaluation include: concluding, criticizing, prioritizing, and recommending.

Capstone Assessment

Assessment of outcomes structured into learning experiences occurring at the end of a program. The experiences involve demonstration of a comprehensive range of program outcomes through some type of product or performance. The outcomes may be those of the major and of the general education program or of the major only. (Palomba & Banta, 1999)

Cohort

A group whose progress is followed by means of measurements at different points in time.

Concept

An abstract, general notion -- a heading that characterizes a set of behaviors and beliefs.

Criteria/Standards

Performance descriptors that indicate how well students will meet expectations of what they should be able to think, know or do. They are descriptive benchmarks against which performance is judged. These criteria or standards may be described in varying gradients of success as in rubrics or in grades. Often they are state in terms of percentages, percentiles or other quantitative measures (Nichols, 2000) (See **Descriptor, Rubrics, Benchmark.**)

Criterion Referenced Tests

A test in which the results can be used to determine a student's progress toward mastery of a content area. Performance is compared to an expected level of mastery in a content area rather than to other students' scores. Such tests usually include questions based on what the student was taught and are designed to measure the student's mastery of designated objectives of an instructional program. The "criterion" is the standard of performance established as the passing score for the test. Scores have meaning in terms of what the student knows or can do, rather than how the test-taker compares to a reference or norm group. Criterion referenced tests can have norms, but comparison to a norm is not the purpose of the assessment.

Criterion referenced tests have also been used to provide information for program evaluation, especially to track the success or progress of programs and student populations that have been involved in change or that are at risk of inequity. In this case, the tests are used to give feedback on progress of groups and individuals.

Curriculum Alignment

The degree to which a curriculum's scope and sequence matches a testing program's evaluation measures, thus ensuring that teachers will use successful completion of the test as a goal of classroom instruction.

Curriculum-embedded or Learning-embedded Assessment

Assessment that occurs simultaneously with learning such as projects, portfolios and "exhibitions." Occurs in the classroom setting, and, if properly designed, students should not be able to tell whether they are being taught or assessed. Tasks or tests are developed from the curriculum or instructional materials.

Cut Score

Score used to determine the minimum performance level needed to pass a competency test. (See **Descriptor** for another type of determiner.)

Descriptor

A set of signs used as a scale against which a performance or product is placed in an evaluation.

Descriptors allow assessment to include clear guidelines for what is and is not valued in student work.

Dimension

Aspects or categories in which performance in a domain or subject area will be judged. Separate descriptors or scoring methods may apply to each dimension of the student's performance assessment.

Direct Assessment Methods

These methods involve students' display of knowledge and skills (e.g. text results, written assignments, presentations, classroom assignments) resulting from learning experience in the class/program. (Palomba & Banta, 1999)

Essay Test

A test that requires students to answer questions in writing. Responses can be brief or extensive. Tests for recall, ability to apply knowledge of a subject to questions about the subject, rather than ability to choose the least incorrect answer from a menu of options.

Evaluation

Both qualitative and quantitative descriptions of student behavior plus value judgments concerning the desirability of that behavior. Using collected information (assessments) to make informed decisions about continued instruction, programs, activities. Exemplar Model of excellence.

Decisions made about assessment findings; deciding about the value of programs/program outcomes; may involve recommendations for changes. (See **Benchmark, Norm, Rubric, Standard.**)

Formative Assessment

Observations which allow one to determine the degree to which students know or are able to do a given learning task, and which identifies the part of the task that the student does not know or is unable to do. Outcomes suggest future steps for teaching and learning.

Assessment conducted during a performance/course/program with the purpose of providing feedback that can be used to modify, shape, and improve a performance/course/program. (Palomba & Banta, 1999) (See **Summative Assessment.**)

Holistic Method/Holistic Scoring

In assessment, assigning a single score based on an overall assessment of performance rather than by scoring or analyzing dimensions individually. The product is considered to be more than the sum of its parts and so the quality of a final product or performance is evaluated rather than the process or dimension of performance. A holistic scoring rubric might combine a number of elements on a single scale. Focused holistic scoring may be used to evaluate a limited portion of a learner's performance.

A type of grading in which an assignment is given an overall score. Possible scores are described in a rating scale. A high score indicates achievement of all aspects of the assignment, while a low score means few if any of the desired outcomes have been achieved. The score levels need to be specific

enough to reveal meaningful, diagnostic information when the scores are aggregated. (Ewell, 1991; Palomba & Banta, 1999).

Indirect Assessment Methods

Assessment methods that involve perceptions of learning rather than actual demonstrations of outcome achievement (e.g. alumni surveys, employer surveys, exit interviews).

Institutional Effectiveness

The measure of what an institution actually achieves.

Item Analysis

Analyzing each item on a test to determine the proportions of students selecting each answer. Can be used to evaluate student strengths and weaknesses; may point to problems with the test's validity and to possible bias.

Journals

Students' personal records and reactions to various aspects of learning and developing ideas. A reflective process often found to consolidate and enhance learning.

Mean

One of several ways of representing a group with a single, typical score. It is figured by adding up all the individual scores in a group and dividing them by the number of people in the group. Can be affected by extremely low or high scores.

Measurement

Quantitative description of student learning and qualitative description of student attitude.

Median

The point on a scale that divides a group into two equal subgroups. Another way to represent a group's scores with a single, typical score. The median is not affected by low or high scores as is the mean. (See **Norm.**)

Metacognition

The knowledge of one's own thinking processes and strategies, and the ability to consciously reflect and act on the knowledge of cognition to modify those processes and strategies.

Mission

A holistic vision of the values and philosophy of a department, program, unit or institution. General education learning goals are often found in the institution's mission statement. (Palomba & Banta, 1999; Allen, 2004)

Modifications

Recommended actions or changes for improving student learning, service delivery, etc. that respond to the respective measurement evaluation.

Multidimensional Assessment

Assessment that gathers information about a broad spectrum of abilities and skills.

Multiple Choice Tests

A test in which students are presented with a question or an incomplete sentence or idea. The students are expected to choose the correct or best answer/completion from a menu of alternatives.

Norm

A distribution of scores obtained from a norm group. The norm is the midpoint (or median) of scores or performance of the students in that group. Fifty percent will score above and fifty percent below the norm.

Norm Group

A random group of students selected by a test developer to take a test to provide a range of scores and establish the percentiles of performance for use in establishing scoring standards.

Norm Referenced Tests

A test in which a student or a group's performance is compared to that of a norm group. The student or group scores will not fall evenly on either side of the median established by the original test takers. The results are relative to the performance of an external group and are designed to be compared with the norm group providing a performance standard. Often used to measure and compare students, schools, districts, and states on the basis of norm-established scales of achievement.

Objectives

Synonymous with outcomes. Statements that describe measurable expectations of what students should be able to think, know or do when they've completed a given educational program. Each statement should describe one expectation; should not bundle several into one statement. The statements must be clear and easily understood by all faculty in the area/department. (See **Outcomes**)

Objective Test

A test for which the scoring procedure is completely specified enabling agreement among different scorers. A correct-answer test.

On-Demand Assessment

An assessment process that takes place as a scheduled event outside the normal routine. An attempt to summarize what students have learned that is not embedded in classroom activity.

Outcomes

An operationally defined educational goal, usually a culminating activity, product, or performance that can be measured. (See **Objectives**)

Percentile

A ranking scale ranging from a low of 1 to a high of 99 with 50 as the median score. A percentile rank indicates the percentage of a reference or norm group obtaining scores equal to or less than the test-taker's score. A percentile score does not refer to the percentage of questions answered correctly, it indicates the test-taker's standing relative to the norm group standard.

Performance-Based Assessment

Direct, systematic observation and rating of student performance of an educational objective, often an ongoing observation over a period of time, and typically involving the creation of products. The assessment may be a continuing interaction between faculty and student and should ideally be part of the learning process. The assessment should be a real-world performance with relevance to the student and learning community. Assessment of the performance is done using a rubric, or analytic scoring guide to aid in objectivity. Performance-based assessment is a test of the ability to apply

knowledge in a real-life setting. Performance of exemplary tasks in the demonstration of intellectual ability.

Evaluation of the product of a learning experience can also be used to evaluate the effectiveness of teaching methods.

Performance Criteria

The standards by which student performance is evaluated. Performance criteria help assessors maintain objectivity and provide students with important information about expectations, giving them a target or goal to strive for.

Portfolio

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. The collection should involve the student in selection of its contents, and should include information about the performance criteria, the rubric or criteria for judging merit, and evidence of student self-reflection or evaluation. It should include representative work, providing a documentation of the learner's performance and a basis for evaluation of the student's progress. Portfolios may include a variety of demonstrations of learning and have been gathered in the form of a physical collection of materials, videos, CD-ROMs, reflective journals, etc.

Portfolio Assessment

A type of direct measure, a performance measure, in which students' assignments are carefully reviewed for evidence of desired learning outcomes. The portfolios contain work selected over a period of time, with materials added as the student progresses through the course/program. In addition, the portfolios usually include students' reflective learning/outcome analysis.

Portfolios may be assessed in a variety of ways. Each piece may be individually scored, or the portfolio might be assessed merely for the presence of required pieces, or a holistic scoring process might be used and an evaluation made on the basis of an overall impression of the student's collected work. It is common that assessors work together to establish consensus of standards or to ensure greater reliability in evaluation of student work. Established criteria are often used by reviewers and students

involved in the process of evaluating progress and achievement of objectives.

Primary Trait Method

Factors or traits (assignment specific) that are considered in scoring an assignment generally stated in a hierarchical scale of three to five incremental levels of achievement quality. For each level on the scale, there is a specific statement that describes expected behavior (criterion) at that level. (Palomba & Banta, 1999; Walvoord & Anderson, 1998).

A type of rubric scoring constructed to assess a specific trait, skill, behavior, or format, or the evaluation of the primary impact of a learning process on a designated audience.

Process

A generalizable method of doing something, generally involving steps or operations which are usually ordered and/or interdependent. Process can be evaluated as part of an assessment, as in the example of evaluating a student's performance during prewriting exercises leading up to the final production of an essay or paper.

Product

The tangible and stable result of a performance or task. An assessment is made of student performance based on evaluation of the product of a demonstration of learning.

Profile

A graphic compilation of the performance of an individual on a series of assessments.

Project

A complex assignment involving more than one type of activity and production. Projects can take a variety of forms, some examples are a mural construction, a shared service project, or other collaborative or individual effort.

Quantitative Methods of Assessment

Methods that rely on numerical scores or ratings. Examples: Surveys, Inventories, Institutional/departmental data, departmental/course-level exams (locally constructed, standardized, etc.).

Qualitative Methods of Assessment

Methods that rely on descriptions rather than numbers. Examples: Ethnographic field studies, logs, journals, participant observation, and open-ended questions on interviews and surveys.

Quartile

The breakdown of an aggregate of percentile rankings into four categories: the 0-25th percentile, 26-50th percentile, etc.

Quintile

The breakdown of an aggregate of percentile rankings into five categories: the 0-20th percentile, 21-40th percentile, etc.

Rating Scale

A scale based on descriptive words or phrases that indicate performance levels. Qualities of a performance are described (e.g., advanced, intermediate, novice) in order to designate a level of achievement. The scale may be used with rubrics or descriptions of each level of performance.

Reliability

The measure of consistency for an assessment instrument. The instrument should yield similar results over time with similar populations in similar circumstances.

Rubric

Some of the definitions of rubric are contradictory. In general a rubric is a scoring guide used in subjective assessments. A rubric implies that a rule defining the criteria of an assessment system is followed in evaluation. A rubric can be an explicit description of performance characteristics corresponding to a point on a rating scale. A scoring rubric makes explicit expected qualities of performance on a rating scale or the definition of a single scoring point on a scale.

A kind of holistic or primary trait scoring in which detailed criteria are delineated and used to discriminate among levels of achievement in assignments, performances, or products.

Sampling

A way to obtain information about a large group by examining a smaller, randomly chosen selection (the sample) of group members. If the sampling is conducted correctly, the results will be representative of the group as a whole. Sampling may also refer to the choice of smaller tasks or processes that will be valid for making inferences about the student's performance in a larger domain. "Matrix sampling" asks different groups to take small segments of a test; the results will reflect the ability of the larger group on a complete range of tasks.

Scale

A classification tool or counting system designed to indicate and measure the degree to which an event or behavior has occurred.

Scale Scores

Scores based on a scale ranging from 001 to 999. Scale scores are useful in comparing performance in one subject area across classes, programs and other large populations, especially in monitoring change over time.

Score

A rating of performance based on a scale or classification.

Scoring Criteria

Rules for assigning a score or the dimensions of proficiency in performance used to describe a student's response to a task. May include rating scales, checklists, answer keys, and other scoring tools. In a subjective assessment situation, a rubric.

Scoring

A package of guidelines intended for people scoring performance assessments. May include instructions for raters, notes on training raters, rating scales, samples of student work exemplifying various levels of performance.

Self-Assessment

A process in which a student engages in a systematic review of a performance, usually for the purpose of improving future performance. May involve comparison with a standard, established criteria. May involve critiquing one's own work or may be a simple description of the performance. Reflection, self-evaluation, metacognition, are related terms.

Standards

Agreed upon values used to measure the quality of student performance, instructional methods, curriculum, etc.

Subjective Test

A test in which the impression or opinion of the assessor determines the score or evaluation of performance. A test in which the answers cannot be known or prescribed in advance.

Summative Assessment

Assessment conducted after a program has been implemented and completed to make judgments about its quality or worth compared to previously defined standards. (Palomba & Banta, 1999)

Evaluation at the conclusion of a unit or units of instruction or an activity or plan to determine or judge student skills and knowledge or effectiveness of a plan or activity. Outcomes are the culmination of a teaching/learning process for a unit, subject, or year's study. (See **Formative Assessment**.)

Triangulation

Multiple lines of evidence pointing to the same conclusion.

Validity

The test measures the desired performance and appropriate inferences can be drawn from the results. The assessment accurately reflects the learning it was designed to measure.

Resources

Allen, Mary. "Articulating Learning Objectives". Workshop presented at WASC/AAHE Collaborative Workshop on Building Learner-Centered Institutions-Developing Institutional Strategies for Assessing and Improving Student Learning. March 2004.

Association for Supervision and Curriculum Development (ASCD)
<http://www.ascd.org/portal/site/ascd/menuitem.5a47c86b3b7b44128716b710e3108a0c/>

Angelo, T. (1995). Improving Classroom Assessment to Improve Learning. Assessment Update, 7(6), 1-2, 13-14.

California State University Sacramento “Assessment Glossary of Terms” 23 August 2004.

<http://www.csus.edu/acad/Portfolios/GE/glossary.htm>

Ewell, P.T. (1991). To Capture the Ineffable: New Forms of Assessment in Higher Education. In G. Grant (ed.), Review of research in Education, no. 17, Washington, DC: American Educational Research Association.

Lyons, N. “Portfolios and Their Consequences: Developing as a Reflective Practitioner” In N. Lyons (ed.), With Portfolio in Hand: Validating the New teacher Professionalism, New York: Teacher’s College Press.

New Horizons for Learning <http://www.newhorizons.org>

Nichols J & Nichols K. (2000). The Departmental Guide and Record Book for Student Outcomes Assessment and Institutional Effectiveness. NY: Agathon Press.

Palomba, C & Banta T. (1999). Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education. San Francisco: Jossey Bass

Walvoord, B. and Anderson V.J., (1998) Effective Grading: A Tool for Learning and Assessment. San Francisco: Jossey Bass.