

Simmons College of Kentucky

**Glossary of Assessment Terminology**

**Accountability:** set of initiatives others take to monitor the results of our actions, and to penalize or reward us based on the outcomes.

**Administrative outcomes:** operational and specific statements derived from a unit's core functions that describe the desired quality of key services within an administrative unit and define exactly what the services should promote.

**Administrative unit operations:** refers to the assessments based on objectives within administrative units that enhance areas of the university in support of student programs and services.

**Administrative unit strategic goal:** broad and generalized statement of action that assists in meeting the mission of the administrative unit and university. Often refers to a long-term time frame.

**Administrative unit strategic objective:** specific statement referring to a short-term time frame and that aligns to the goal.

**Administrative unit strategic outcome:** describes a change in students that results from a provided learning experience.

**Alignment:** process of assuring that learning outcomes, curriculum and instruction, and assessment all support and match each other. (The Higher Education Academy)

**anchors:** samples of student work collected to provide examples that indicate different levels from a scoring rubric.

**Annual update:** A brief report from each academic program based on its assessment plan and submitted annually, which outlines how evidence was used to improve student learning outcomes through curricular and/or other changes or to document that no changes were needed.

**Archival/Peer records:** Biographical, academic, or other file data available from the college or other agencies and institutions.

**Assessment:** (1) A method for analyzing and describing student learning outcomes or program achievement of objectives. Many assessments are not tests. For students, a reading miscue analysis is an assessment, a direct observation of student behavior can be an assessment, and a student conference can be an assessment. For programs, a senior exit interview can be an assessment, and an employer survey of satisfaction with graduates can be an assessment. Good assessment requires feedback to those who are being assessed so that they can use that information to make improvements. A good assessment program requires using a variety of assessment instruments each one designed to discover unique aspects of student learning outcomes and achievement of program objectives.

**Assessment:** (2) systematic collection, review, and use of information about educational programs undertaken for the purpose of improving student learning and development.  
(Marchese)

**Assessment for accountability:** Assessment of some unit (could be a program, department, college or entire institution) to satisfy stakeholders external to the unit itself. Results are summative and are often compared across units. For 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.

**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 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.

**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 plan:** A document that outlines the student learning outcomes and program objectives, the direct and indirect assessment methods used to demonstrate the attainment of each outcome/objective, a brief explanation of the assessment methods, an indication of which outcome(s)/objectives is/are addressed by each method, the intervals at which evidence is collected and reviewed, and the individual(s) responsible for the collection/review of evidence.

**Assessment system:** comprehensive and integrated set of assessment measures that provides information for use in monitoring student learning outcomes and managing and improving academic programs, student development, and administrative unit operations to promote continuous improvement, enhance institutional effectiveness, and ensure accountability.

**Authentic assessment(s) (1):** real-world activities that professionals in the discipline may encounter. Assessment can be conducted at fieldwork sites in which students work with clients or address problems. (Allen)

**Authentic assessment (2):** An assessment that measures a student's ability to perform a —real world task in the way professionals in the field would perform it. An authentic writing task might arise if students had been reading about nutrition and decided to ask the school to provide healthy snacks rather than candy machines; their writing would be assessed in terms of the response it received from the principal and/or school board. An authentic reading task would require assessing a student's understanding of a book he or she had selected to read without any suggestions or restrictions by the teacher. Opportunities for truly authentic assessment do not occur regularly in most classrooms.

**Authentic performance assessment:** Since regular opportunities for truly authentic tasks come infrequently in most classrooms, this term generally indicates an evaluation of a student's ability to perform a complex task that is common in the classroom. An authentic performance assessment in a science class would occur when a student is asked to perform an experiment and write a lab report; an authentic writing performance assessment would occur when a student generated a topic, created multiple drafts, sought outside opinions and editorial assistance, and published his or her paper in a classroom magazine or web page. Taking a test over science terms or labeling the parts of a sentence would not be authentic performance assessment. Writing an essay in a limited amount of time in response to a prompt is not an authentic writing assessment either because these circumstances do not match the way writing is usually produced outside of school.

**Backload (--ed, --ing):** Amount of effort required after the data collection.

**Behavioral observations:** Measuring the frequency, duration, topology, etc. of student actions, usually in a natural setting with non-interactive methods, for example, formal or informal observations of a classroom. Observations are most often made by an individual and can be augmented by audio or videotape.

**Benchmarking:** the process of comparing institutions' information and assessment results with other institutions, often their peers. (Suskie)

**Competency (1):** Level at which performance is acceptable.

**Competency (2):** A group of characteristics, native or acquired, which indicate an individual's ability to acquire skills in a given area.

**Confounded:** The situation in which the effect of a controlled variable is inextricably mixed with that of another, uncontrolled variable.

**Convergent validity:** General agreement among ratings, gathered independently of one another, where measures should be theoretically related.

**Commercial, norm-referenced, standardized exams:** Group administered, mostly or entirely multiple-choice, "objective" tests in one or more curricular areas. Scores are based on comparison with a reference or norm group. Typically must be purchased from a private vendor.

**Constructed-response:** assessment method that requires students to construct a tangible product or perform a demonstration to show what they know and are able to do.

**Course embedded assessments (1):** assessments generated from assignments already in place in the classroom. (Palomba & Banta)

**Course-embedded assessment (2):** Course-embedded assessment refers to techniques that can be utilized within the context of a classroom (one class period, several or over the duration of the course) to assess students' learning, as individuals and in groups. When used in conjunction with other assessment tools, course embedded assessment can provide valuable information at specific points of a program.

For example, faculty members teaching multiple sections of an introductory course might include a common pre-test to determine student knowledge, skills and dispositions in a particular field at program admission. There are literally hundreds of classroom assessment techniques, limited only by the instructor's imagination (see also embedded assessment).

**Course objectives:** similar to goals but express the intended content to be covered in a course. They are used to describe specific behaviors that the student should exhibit. (Palomba & Banta)

**Criterion-referenced:** Criterion-referenced tests determine what test-takers can do and what they know, not how they compare to others. Criterion-referenced tests report on how well students are doing relative to a predetermined performance level on a specified set of educational goals or outcomes included in the curriculum. For example, student scores on tests as indicators of student performance on standardized exams.

**Curriculum mapping:** matrix used to indicate where student learning outcomes are covered in each course. Level of instructional emphasis or assessment of where the student learning outcome takes place may also be indicated.

**Direct measures:** assessment that requires students to demonstrate their achievement directly from their work. (Allen)

**Embedded assessment:** A means of gathering information about student learning that is built into and a natural part of the teaching learning process. Often used for assessment purposes in 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).

**E-portfolio (electronic portfolio):** An electronic format of a collection of work developed across varied contexts over time. The e-portfolio can advance learning by providing students and/or faculty with a way to organize, archive and display pieces of work. The electronic format allows faculty and other professionals to evaluate student portfolios using technology, which may include the Internet, CD-ROM, video, animation or audio. Electronic portfolios are becoming a popular alternative to traditional paper-based portfolios because they offer practitioners and peers the opportunity to review, communicate and assess portfolios in an asynchronous manner (see also portfolios also called course-embedded assessment).

**Evaluation (1):** Depending on the context, evaluation may mean either assessment or test. Many test manufacturers and teachers use these three terms interchangeably which means you have to pay close attention to how the terms are being used and why they are being used that way. For instance, tests that do not provide any immediate, helpful feedback to students and teachers should never be called —assessments, but many testing companies and some administrators use this term to describe tests that return only score numbers to students and/or teachers (Palomba & Banta).

**Evaluation (2):** When used for most educational settings, evaluation means to measure, compare, and judge the quality of student work, schools, or specific educational programs.

**Evaluation (3):** A value judgment about the results of assessment data. For example, evaluation of student learning requires that educators compare student performance to a standard to determine how the student measures up. Depending on the result, decisions are made regarding whether and how to improve student performance.

**Exit and other interviews:** Asking individuals to share their perceptions of their own attitudes and/or behaviors or those of others, evaluating student reports of their attitudes and/or behaviors in a face-to-face-dialogue.

**External Assessment:** Use of criteria (rubric) or an instrument developed by an individual or organization external to the one being assessed.

**External examiner:** Using an expert in the field from outside your program, usually from a similar program at another institution to conduct, evaluate, or supplement assessment of your students. Information can be obtained from external evaluators using many methods including surveys, interviews, etc.

**External validity:** External validity refers to the extent to which the results of a study are generalizable or transferable to other settings. Generalizability is the extent to which assessment findings and conclusions from a study conducted on a sample population can be applied to the population at large. Transferability is the ability to apply the findings in one context to another similar context.

**Fairness (1):** Assessment or test that provides an even playing field for all students. Absolute fairness is an impossible goal because all tests privilege some test takers over others; standardized tests provide one kind of fairness while performance tests provide another. The highest degree of fairness can be achieved when students can demonstrate their understanding in a variety of ways.

**Fairness (2):** Teachers, students, parents and administrators agree that the instrument has validity, reliability, and authenticity, and they therefore have confidence in the instrument and its results.

**Focus groups:** Typically conducted with 7-12 individuals who share certain characteristics that are related to a particular topic, area or assessment question. Group discussions are conducted by a trained moderator with participants to identify trends/patterns in perceptions. The moderator's purpose is to provide direction and set the tone for the group discussion, encourage active participation from all group members, and manage time. Moderators must not allow their own biases to enter, verbally or nonverbally. Careful and systematic analysis of the discussions provides information that can be used to assess and/or improve the desired outcome.

**Follow-up report:** A report requested by the Academic Planning Council (APC) following program review to address specific issue(s)/concern(s) that result from the Council's examination review of program review documents. The report is submitted within the time frame identified by the Council prior to the program's full review by the APC.

**Forced-choice:** The respondent only has a choice among given responses (e.g., very poor, poor, fair, good, very good).

**Formative assessment (1):** assessing student learning over time; provides valuable information about how well students are progressing towards an institution's or program's expectations. (Maki)

**Formative assessment (2):** 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. Assessment feedback is short term in duration. 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.

**Frontload (--ed, --ing):** Amount of effort required in the early stage of assessment method development or data collection.

**Generalization (generalizability):** The extent to which assessment findings and conclusions from a study conducted on a sample population can be applied to the population at large.

**Goal-free evaluation:** Goal-free evaluation focuses on actual outcomes rather than intended program outcomes. Evaluation is done without prior knowledge of the goals of the program.

**High stakes test:** A test whose results have important, direct consequences for examinees, program, or institutions tested.

**“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.

**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.

**Indirect measures:** assessments of student learning that are based on opinion, often the students. (Allen)

**Institutional effectiveness:** documented process of measuring how well an institution is achieving its mission and addressing its strategic plan for the purpose of continuous improvement of student learning, student development, and administrative unit operations.

**Institutional portfolios:** Institutional portfolios provide a means of assessing the impact of the entire educational experience on student learning. They can be used to drive internal improvement and external accountability. Like student portfolios, they allow for internal improvement and external accountability, but on the level of the whole institution (see also portfolios).

**Inter-rater reliability:** The degree to which different raters/observers give consistent estimates of the same phenomenon.

**Internal validity:** Internal validity refers to (1) the rigor with which the study was conducted (e.g., the study's design, the care taken to conduct measurements, and decisions concerning what was and wasn't measured) and (2) the extent to which the designers of a study have taken into account alternative explanations for any causal relationships they explore.

**Local assessment:** Means and methods that are developed by an institution's faculty based on their teaching approaches, students, and learning goals. Is an antonym for —external assessment. 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.

**Locally developed exams:** Objective and/or subjective tests designed by faculty of the program or course sequence being evaluated.

**Longitudinal studies:** Data collected from the same population at different points in time.

**Metric:** what is being assessed.

**Norm (--ative):** A performance standard that is established by a reference group and that describes average or typical performance. Usually norms are determined by testing a representative group and then calculating the group's test performance.

**Norm-reference:** A norm-referenced test is one designed to highlight achievement differences between and among students to produce a dependable rank order of students across a continuum of achievement from high achievers to low achievers.

**Objective:** planned or intended outcome.

**Observer effect:** The degree to which the assessment results are affected by the presence of an observer.

**Open-ended:** Assessment questions that are designed to permit spontaneous and unguided responses.

**Operational (--ize):** Defining a term or object so that it can be measured. Generally states the operations or procedures used that distinguish it from others.

**Oral examination :** An assessment of student knowledge levels through a face-to-face dialogue between the student and examiner-usually faculty.

**Performance appraisals:** A competency-based method whereby abilities are measured in most direct, real-world approach. Systematic measurement of overt demonstration of acquired skills.

**Performance assessment (1):** A method for assessing how well students use their knowledge and skills in order to do something. Music students performing a new piece of music before a panel of judges are undergoing performance assessment; students who are expected to demonstrate an understanding of basic grammar, spelling, and organizational skills while writing a paper are undergoing performance assessment; business students asked to write a proposal to solve a problem presented in a case study are undergoing performance assessment.

**Performance assessment (2):** process of using student activities or products, as opposed to tests or surveys, to evaluate students' knowledge, skills, and development. (Palomba & Banta)

**Performance criteria:** can be defined in terms of —learning outcomes statements which describe, using action verbs, student learning or behavior rather than teacher behavior; and describe an intended outcome rather than subject matter coverage. (Palomba & Banta)

**Portfolios:** Collections of multiple student work samples usually compiled over time and rated using rubrics. The design of a portfolio is dependent upon how the scoring results are going to be used.

**Program assessment:** does not focus on an individual student. Rather, the emphasis is on what and how an educational program is contributing to the learning, growth and development of students as a group. Goals are broad statements that describe the long-term program targets or (University of Central Florida)

**Program goals:** broad statements that describe the long-term program targets or directions of development. Stated in broad terms what the program wants to accomplish (in student learning outcomes) or desires to become over the next several years. (University of Central Florida)

**Program review:** The administrative and peer review of academic programs conducted on an five to eight-year cycle, the results of which are reported to the Cabinet and the Board of Trustees. This review includes a comprehensive analysis of the structure, processes, and outcomes of the program. The outcomes reported in the program reviews include program outcomes (e.g. costs, degrees awarded) as well as student learning outcomes (i.e. what students know and can do at the completion of the program)

**Qualitative methods of assessment:** Methods that rely on descriptions rather than numbers. Examples: ethnographic field studies, logs, journals, participant observations, open-ended questions on interviews and surveys.

**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.)

**Reliability:** The extent to which an experiment, test or any measuring procedure yields the same result on repeated trials.

**Rubric:** scoring tool that provides the specific expectations for an assignment. Rubrics divide an assignment into the critical elements to be examined and provide detailed descriptions of what constitutes acceptable or unacceptable levels of performance for each of those elements.

1. Holistic: rubric that measures the overall quality of an artifact, performance, or portfolio. (Krajcik, Czerniak, & Berger)

2. Analytic: rubric where criteria are broken down into critical elements, content/coverage, of a performance.

**Salience:** A striking point or feature.

**Selected-response:** assessment method that requires students to select a response from a provided list or supply a brief answer. Examples: multiple choice, true/false, matching, or essay tests.

**Simulations:** A competency-based measure where a person's abilities are measured in a situation that approximates a "real world" setting. Simulation is primarily used when it is impractical to observe a person performing a task in a real world situation (e.g. on the job).

**Stakeholder:** Anyone who has a vested interest in the outcome of the program/project. In a high stakes standardized test (a graduation requirement, for example), when students' scores are aggregated and published in the paper by school, the stakeholders include students, teachers, parents, school and district administrators, lawmakers (including the governor), and even real estate agents. It is always interesting to note which stakeholders seem to have the most at risk and which stakeholders seem to have the most power; these groups are seldom the same.

**Standard:** The performance level associated with a particular rating or grade on a test. For instance, 90% may be the standard for an A in a particular course; on a standardized test, a cutting score or cut point is used to determine the difference between one standard and the next.

**Standard-based assessment:** A standard-based assessment assesses learner achievement in relation to set standards.

**Standardized test (1):** This kind of test (sometimes called —norm-referenced) is used to measure the performance of a group against that of a larger group. Standardized tests are often

used in large-scale assessment projects, where the overall results of the group are more important than specific data on each individual client. Standardized tests are not authentic. They are most useful for reporting summative information, and are least useful for classroom diagnosis and formative purposes.

**Standardized test (2):** assessment where conditions of administration and scoring are constant. A well-designed standardized test will have a set of procedures for administration that can be implemented by all users. A standard set of introductory comments and directions are developed and used by all test takers. (Palomba & Banta)

**Standards:** Widely recognized models of excellence; term commonly used to describe achievement goals. Standards are always prescriptive because they tell us what “should be”

**Status report:** A description of the implementation of the plan's assessment methods, the findings (evidence) from assessment methods, how the findings were used in decisions to maintain or improve student learning (academic programs) or unit outcomes (support units), the results of previous changes to improve outcomes, and the need for additional information and/or resources to implement an approved assessment plan or gather additional evidence.

**Student development:** refers to the assessments within our division of student affairs and other administrative units that promote out-of-class student learning, growth, and development outcomes through structured programs and services.

**Student learning:** refers to the measurable outcomes of what students should know and are able to do as a result of their course work and educational experiences at our institution.

**Student learning outcomes:** measurable statements of what students should know and be able to do as a result of their course work and educational experiences at an institution or in a program of study. (Maki)

**Summative assessment:** assessment of student learning at the end of a program or course of study; provides information about patterns of student achievement without institutional or programmatic opportunity to improve students' achievement and without student opportunity to reflect on how to improve and demonstrate that improvement. (Maki)

**Taxonomic schemes:** a hierarchical structure of data arranged in a classification system.

**Test:** A formal assessment of student achievement. Teacher made tests can take many forms; external tests are always standardized. A portfolio can be used as a test, as can a project or exhibition.

**Third party:** Person(s) other than those directly involved in the educational process (e.g., employers, parents, consultants).

**Topology:** Mapping of the relationships among subjects.

**Triangulate (triangulation) (1):** The use of a combination of assessment methods in a study. An example of triangulation would be an assessment that incorporated surveys, interviews, and observations.

**Triangulation (2):** collection of data from multiple measures in order to show consistency of results. (Allen)

**Utility (1):** Usefulness of assessment results.

**Utility (2):** The relative value of an outcome with respect to a set of other possible outcomes. Hence test utility refers to an evaluation, often in cost-benefit form, of the relative value of using a test vs. not using it, of using a test in one manner vs. another, or of using one test vs. another test.

**Validity:** Validity refers to the degree to which a study accurately reflects or assesses the specific concept that the researcher is attempting to measure. Validity has three components: *relevance* - the option measures your educational objective as directly as possible *accuracy* - the option measures your educational objective as precisely as possible *utility* - the option provides formative and summative results with clear implications for educational program evaluation and improvement

**Value-added assessment strategy:** assessment perspective that gathers longitudinal data, both quantitative and qualitative information, with the intent of examining the impact of the university on student learning. (Suskie)

**Variable (variability):** Observable characteristics that vary among individuals responses.

**Written surveys/questionnaires:** Asking individuals to share their perceptions about the study target-e.g. their own or others skills/attitudes/behavior, or program/course qualities and attributes.

## **References**

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

Frye, S. (1999). *Assessment, accountability, and student learning outcomes*.

Retrieved May 20, 2008, for the Western Washington University Web Site:

<http://www.ac.wvu.edu/~dialogue/issue2.html>

Krajcik, J., Czerniak, C., & Berger, C. (2003). *Teaching children science: A project based approach*. Retrieved May 20, 2008, from the McGraw Hill: Online

Learning Center Web Site:

[http://highered.mcgraw-hill.com/sites/0072486740/student\\_view0/key\\_terms.html](http://highered.mcgraw-hill.com/sites/0072486740/student_view0/key_terms.html)

Maki, P. (2004). *Assessing for learning: Building a sustainable commitment across the*

institution. Sterling, VA: Sterling.

Marchese, T. J. (1987). Third down, ten years to go. AAHE Bulletin, 40, 3-8.

Nichols, J. (2005). The administrative unit assessment handbook: Measuring student support services and administrative objectives. Retrieved February 3, 2008 from the University of Central Florida Web Site:

[http://iaaweb.ucf.edu/oeas/adm\\_assess\\_handbook.pdf](http://iaaweb.ucf.edu/oeas/adm_assess_handbook.pdf)

Palomba, C. A., & Banta, T. W. (1999). Assessment essentials: Planning, implemented, and improving assessment in higher education. San Francisco. Jossey Bass.

Stevens, D.D. & Levi, A. J. (2005). Introduction to rubrics: An assessment tool to save grading time, convey effective feedback and promote student learning. Sterling, VA: Stylus.

Suskie, L. (2004). Assessing student learning: A common sense guide. San Francisco, CA: Josey-Bass.

The Higher Education Academy (2008). Constructive alignment and why it is important to the learning process. Retrieved May 20, 2008, from The Higher Education Academy Web Site: [http://www.engsc.ac.uk/er/theory/constructive\\_alignment.asp](http://www.engsc.ac.uk/er/theory/constructive_alignment.asp)

University of Central Florida (UCF). Program assessment handbook: Guidelines for planning and implementing quality enhancing efforts of program and student learning outcomes. Retrieved February 3, 2008 from University of Central Florida Web Site: [http://iaaweb.ucf.edu/assessment/support/acad\\_assess\\_handbook.pdf](http://iaaweb.ucf.edu/assessment/support/acad_assess_handbook.pdf)

