

## ASSESSMENT GLOSSARY

This glossary provides definitions of terms as used in this text. For many of the terms, multiple definitions can be found in the literature; also, technical usage may differ from common usage.

---

**Accountability** - The public reporting of student, program, or institutional data for the purpose of program improvement.

**Achievement** - The attainment of new knowledge or understanding as the result of an educational activity.

**Advanced Placement Exams** - Tests and assessments given at the beginning of an instructional program to determine to what extent students have already achieved the *objectives* of the planned instruction, so that students may be placed in more advanced course work.

**Affective** – A person’s dispositions or attitudes or behaviors.

**Angoff Method** – Established in 1971, the Angoff educational measurement method charges judges to assign a probability to each item on a test. Each judge is asked to state the probability that the “minimally acceptable person” would answer each item correctly. The sum of these probabilities, or proportions, would represent the minimally acceptable score.

**Assessment** - The systematic process of determining educational objectives, gathering, analyzing and using information about student learning outcomes to make decisions about programs, individual student progress, and/or accountability.

**Assessment Feedback** - Using assessment results to help guide program change and improvement.

**Assessment (Formative)** - Purposeful, ongoing collection of information regarding how students are learning while there is still the opportunity to make improvements. Both teachers and students use the information to guide continuous improvement toward the intended learning.

**Assessment Literacy** - The term used to describe a set of knowledge and skills educators need to know and be able to use regarding student learning.

**Assessment Methods** - The different ways used by schools to evaluate learning. There are four basic categories of assessment methods: written forced response, essay, performance assessment, and portfolio.

**Assessment (Summative)** - An assessment measure of achievement at the end of an instructional unit, course of study, or program.

**Authentic (Performance) Assessment** - Assessment technique that involves the gathering of data through systematic observation of a behavior, process, or product and the evaluation of those based on a clearly articulated set of performance criteria, which serve as the basis for evaluative judgments.

**Battery** – A set of tests usually administered as a unit. The scores on the different tests in the set usually are scaled so they can readily be compared or used in combination for decision making.

**Benchmark (Standard, Cut-score)** - A criterion-referenced objective. Performance data that are used to set a level for comparison, either between different programs or over time for the same program. If data from another exemplary program are chosen as benchmarks, it becomes a target to strive for, rather than a baseline to improve upon.

**Career and Technical Education** - Organized educational activities that offer a sequence of courses that provide individuals with the academic and technical knowledge and skills needed to prepare for further education and for careers in current or emerging employment sectors; and includes competency-based applied learning that contributes to the academic knowledge, higher-order reasoning and problem-solving skills, work attitudes, general employability skills, technical, and occupation-specific skills of participating students

**Certification** - Verification that a level of achievement has met designated standards.

**Coefficient Alpha** – An internal consistency reliability coefficient based on the number of parts into which the test is portioned (e.g., items, subtest, or raters), the interrelationships of the parts, and the total test score variance. Also called *Cronbach's alpha* and, for dichotomous items, *KR 20*.

**Cognitive Assessment** – The process of systematically gathering test scores and related data in order to make judgments about an individual's ability to perform various mental activities involved in the processing, acquisition, retention, conceptualization, and organization of sensory, perceptual, verbal, spatial, and psychomotor information.

**Competency** - Demonstration of the ability to perform a specific task or achieve specified criteria.

**Competency Listing** - (Also *critical competencies, core competencies, standards, duties and tasks*) Terms used interchangeably throughout the NOCTI web site, referring to the broad categories and particular skills covered on a particular assessment.

**Computer-based Testing** - (Also *online testing*) An assessment administered using a computer online system instead of hard copy/pencil and paper instrument.

**Content Validity** - A test has content validity if the item contents are indicators of the characteristic it is designed to measure.

**Continuous Quality Improvement** - CQI is a methodology intent on meeting or exceeding customer requirements by continuous improvement and innovation in products, processes, and services.

**Course** – A complete body of prescribed studies constituting a curriculum.

**Credentialing** - Method of assuring students, parents/guardians, and the community that students are prepared with the knowledge and skills to be successful as they move from school to school and to their chosen career.

**Criterion-Referenced Assessment** - An assessment that measures student progress toward specific curriculum goals or standards. Scores are reported as comparisons to predefined acceptable levels of performance rather than through comparisons to other students.

**Curriculum** - A collection of statements of the objectives, or learning goals, to be taught in each subject at an individual grade level. Curriculum guides what is taught and how the subject is taught in the classroom.

**Cut-score** - An established score which an individual needs to achieve on an assessment to demonstrate minimal competency.

**DACUM** – An acronym for Developing A Curriculum is a Canadian term. It is an approach to job, occupational, process, and functional analysis that involves bringing a committee of expert workers together under the leadership of a trained facilitator. Modified brainstorming techniques are used to specify in detail the duties and tasks successful workers in the targeted occupation must perform. The general knowledge and skills needed, important worker behaviors, tools and equipment, future trends, and concerns are also identified. The purpose of this comprehensive job and task analysis is to establish content for both curriculum and assessment instruments designed to measure related learner competency.

**Demographics** – Characteristics of human populations.

**Distracter** – An incorrect answer presented as one of the response alternatives in a multiple-choice test.

**Duty** – This is an arbitrary division of a job (or occupational area) consisting of related tasks that are grouped for descriptive purposes. Duties are usually stated as general areas of responsibility, with actions words ending in *ing*.

**Ebel Method** – The Ebel educational measurement method was developed in 1979. Judges rate test items along two dimensions, relevance and difficulty, using four levels of relevance (essential, important, acceptable, and questionable) and three levels of difficulty (easy, medium, and hard). Next, the judges are asked to do two things: locate each of the test items in the proper cell, based upon their relevance and difficulty, and assign a percentage to each cell. The percentage is on the percentage of items in the cell that a minimally qualified examinee should be able to answer. Then the number of test items in each cell is multiplied by the appropriate percentage (agreed upon by the judges), and the sum of all the cells, when divided by the total number of test items, yields the standard or estimated test score for the overall test.

**Employability (Workplace Readiness)** – A person’s ability to be successful in the workplace based on factors such as labor market, occupational information, values clarification and personal understanding, career planning and decision making, and job search techniques, including daily living skills, personal skills, work ethic and initiative, *reliability* and accepting responsibilities involved in maintaining a job.

**Evaluation** - The process of collecting information from multiple sources to make judgments (assign a grade, for example) about how well students have learned and about program effectiveness.

**Evaluator** - An individual who administers and evaluates practical skills on performance-based assessments.

**Evaluator Guide** - A reference and instructional booklet for evaluators, designed to complement performance-based assessments.

**Evaluation Feedback** – Responses or information that tells learners how their assessment scores compare to others or that provides a judgment summarizing the quality of the learning they have accomplished. Letter grades, numbers, symbols, and written phrases are typically used to deliver this type of feedback.

**Exit Surveys** - Information obtained from students on completion of their studies. This includes information about student growth and change, satisfaction with their academic programs, their experiences in their majors, and their immediate and future plans.

**Experienced Worker Assessment** - Assessments for individuals with advanced skills and/or more than three to five years of experience in the occupational field being assessed.

**Fairness (in testing practice)** - This concept encompasses all testing design and practice and points toward the necessity of ensuring that measures and consequences from measures are equitable across different subgroups of examinees.

**Feedback** – The return of information about the result of a process or activity. Feedback is an evaluative response.

**Formative (Evaluation)** - An assessment that is used for making improvements (individual or program level) rather than for making final decisions or for accountability purposes.

**Goals - Statements** describing general expectations regarding student outcomes.

**High Stakes Assessment** - Use of the results of an assessment to determine attainment of a goal such as completing a program of study, receiving certification, or moving to the next level of a program of study.

**Irregularity Report** – Form that NOCTI clients can complete to report difficulties and errors in the implementation of an assessment or make suggestions regarding an assessment. The irregularity Form is also used to report significant occurrences observed while proctoring an assessment that may affect test takers assessment results.

**Item** - (Also *question*) A term used to refer to a question on a written assessment or a job and task in a performance assessment.

**Item Bank and/or Pool** – The aggregate of all items from which a particular test or scale is selected during test development, or the total set of items from which a particular test is selected for a test taker during adaptive testing.

**Jaeger Method** – In 1978, Jaeger offered a method for standard setting on the North Carolina High School Competency Test. His method uses judges from a variety of backgrounds and employs normative data. Rather than ask judges to relate to the term “minimal competence,” the judges are asked to respond to questions such as, “Should every high school graduate be able to answer the item correctly?” “\_Yes, \_, NO.” After a series of processes involving judges from various areas of expertise, and after the presentation of some normative data, standards determined by all groups of judges of the same type are pooled, a median is computed for each type of judge. The minimum median across all groups is selected as the standard. Jaeger’s method has considerable merit for setting standards on very important tests.

**Job** – A specific position requiring the performance of specific duties and related tasks- generally the same tasks are performed by all workers having the same title.

**Job Ready Assessment** - Assessments for program completers of a secondary or post secondary career and technical program at the job entry level.

**Learning Outcomes** - The specific knowledge or skills students actually acquire/develop through their educational experience.

**Likert Scale** - The scale used in items that allow respondents to indicate their level of agreement with a statement by marking their response over a five point scale, usually ranging from strongly agree to strongly disagree.

**Longitudinal Data** - Data collected on the same individuals over time for use in a long-term study.

**Mean** - The average of all the elements (individual's test answers) in a sample. To compute a sample mean, add up all the sample values and divide by the size of the sample.

**Median** - The value on a measure or test for which roughly half the data in a sample are smaller and roughly half the data are larger.

**Meta-analysis** – A statistical method for analyzing research data in which the results from several independent, comparable studies are combined to determine the size of an overall effect or the degree of relationship between two variables.

**Mode** - The most frequently occurring value in a set of data based on a particular measure or test.

**Module** – A type of learning package that usually includes a performance objective, enabling objectives, essential attitudinal and cognitive information as planning or practice activities, and feedback activities. While modules are developed in a wide variety of formats, most are self-contained, transportable, and designed for either individual or group use.

**Multiple Measures** – A number of pieces of assessment data gathered from a variety of sources/tests.

**NCS OP Scan6** - Test scoring technology utilized by NOCTI and The Whitener Group as of 2004.

**Nedelsky Method** – In 1954, Nedelsky's educational measurement method was developed. Under that method, judges are asked to identify distracters in multiple-choice test items that they feel minimally competent students should be able to eliminate as incorrect. The minimum passing level for each item then becomes the reciprocal of the remaining alternatives. The judges proceed with every test item and, upon completion of the judging process each judge sums the minimum passing levels across all test items to obtain a standard. Individual judges' standards are then averaged to obtain a standard for the test.

**Norm-Referenced Assessment** - An interpretation of scores on a test that focuses on the rank ordering of students not their performance in relation to specific criteria.

**Objective Measures** - Tests or measures that can be scored without reliance upon subjective judgments (e.g., tests comprised of multiple-choice items).

**Objectives** – Statements of the specific knowledge, skills, or attitudes that students are expected to achieve through their educational experience.

**Occupational Cluster** – A grouping of jobs that are related by the similarity of the tasks performed by the various workers.

**Occupational Skills Attainment** - Proficiency to perform actual tasks and technical functions required by certain occupational fields at entry, intermediate or advanced levels.

**Online Testing** - (NOCTI's online testing) A program that offers selected assessments delivered and scored through an internet-based computer system.

**Outcomes** - The specific knowledge, skills, or developmental attributes that students actually develop through their educational experience.

**Performance Assessment** - Assessment based on students' performance of selected jobs developed by subject matter experts in the particular field of study being tested.

**Percentile Rank** - The score used most frequently to report individual student results on norm-referenced tests. A score at the 63<sup>rd</sup> percentile means that the student scored as high or better than 63 percent of the students who originally took the test.

**Pilot Study/Assessment** - Often refers to a small, scaled down assessment designed to test the validity of measures and manipulations of a planned full-scale study. A pilot can also refer to the initial administration of new assessment items/procedures with the intent of evaluating and revising the items/procedures for future use.

**Pilot Testing** - The initial administration of a new or revised assessment, with the intent of evaluating and/or adjusting the assessment prior to national, full-scale administration.

**Placement Exams** – Instruments used to determine if a student's qualifications for entering a program or course are at an appropriate level to begin their studies.

**Population** - All possible individuals making up a group of interest in a study or assessment.

**Portfolio** – A collection of a student's work (e.g., products developed or services delivered).

**Post Secondary** – This term refers to education beyond high school - at the college or university level.

**Post-Test (Program)** - An assessment measurement administered after program completion, usually for the purpose of documenting attainment of program-related objectives or comparing the score to an earlier pre-test measure on the same or similar content.

**Pre-Test (Program)** - An assessment measurement administered prior to course initiation, usually for the purpose of identifying existing skills and/or knowledge or for comparing to a post-test measurement of the same or similar content.

**Proctor** – Individual who is knowledgeable regarding administration of objective-type tests and is very detail oriented. Capable of monitoring groups of people and providing succinct instructions. Also assists with training and oversight of Evaluators for NOCTI's Performance Test.

**Proficiency** - Achieving competency on a predetermined standard.

**Program Improvement** - A judgment by experts of the worth of a curriculum, based primarily on observations involving a comparison between student performance indicators and objectives as well as actual student performance on either standardized or evaluator-constructed measures.

**Quality** –The degree to which an educational institution and/or its programs and services are fitted to its purpose, satisfactory to its clients, and of a grade equivalent to other similar institutions.

**Random Sample** - An arbitrary sub-group of persons/items/observations drawn from and meant to represent a larger population.

**Range** - The distance between the highest and lowest score on a given measure/test.

**Rater** - A person who evaluates or judges student performance on an assessment against specific criteria.

**Rating Scale** - A series of items or statements that describe the degree or extent of an aspect of a skill or a personal trait.

**Raw Score** - A measure prior to scaling.

**Reliability** – The degree to which a test consistently measures what it is intended to measure.

**Research** – An approach/method for acquiring knowledge and identifying the causes for behavior.

**Results or Findings** - The outcomes of an assessment study or the experiment.

**Rubric** - A set of scoring guidelines that can be used by raters to evaluate student's work.

**Sample** – A selected subset of entities called sampling units (e.g., test takers, items, etc.) from a larger specified set of such entities, called the population.

**Sample Assessment** - Several sample questions from the written test and a sample job from the performance test. The sample assessments on the NOCTI web site provide summarized breakdowns of the contents covered by NOCTI's different occupational assessments/tests. The sample assessments supply a list of the competencies for each test with a summary of the skill areas covered, several sample questions from the written test, a sample job from the performance test, and a percentage breakdown of the test contents.

**Sampling Error** - The deviation between characteristics of a sample and the related characteristics of a population.

**Satisfaction** - A student's attitude towards their educational environment.

**SCANS Skills** - Department of Labor Commission taxonomy of skills and attributes needed for effective job performance and a set of criteria by which to measure their acquisition; in addition to technical skills, includes five competencies (resources, interpersonal, information, systems, technology) and a three part set of foundation skills (basic skills, thinking skills, personal qualities).

**Scoring Rubric** – The established criteria, including rules, principles, and illustrations, used in scoring individual items and clusters of items. The term usually refers to the scoring procedures for assessment tasks that do not provide enumerated responses from which test takers make a choice. Scoring rubrics vary in the degree of judgment entailed, in the number of distinct score levels defined, in the latitude given scorers for assigning intermediate or fractional score values, and in other ways.

**Secondary** – This term refers to education at the high school level.

**Significance** - The likelihood that relationships observed in a sample or between samples can not be attributed to sampling error alone.

**Skills** - Observable behaviors that document levels of competence (i.e., knowledge, comprehension, application, analysis, synthesis, and evaluation).

**Standards** - The broadest of a family of terms referring to statements of expectations for student learning, including content standards, performance standards, and benchmarks.

**Standard Deviation** - The square root of the variance, used as a measure of dispersion or spread among a group of scores.

**Standard Error of the Measurement** - The standard error is the standard deviation of the sampling distribution of a statistic.

**Stem** – The part of an objective test question or item that poses the question to be answered or the problem to be solved.

**Student Outcome** - The achieved results or the actual consequences of what a student has demonstrated or accomplished; including academic and occupational, as well as the intellectual, personal, civic development, attitudes, values, and beliefs that students attain as a result of their educational experience.

**Subgroup** – A group that is a subset of a larger group.

**Subjective Measures** - Tests or measures that are scored or rated using some degree of subjective judgment (e.g., performance).

**Task** – A meaningful unit of work activity generally performed on the job by one worker within a limited period of time. A task must be logical and necessary to achieving a single objective or output.

**Team based** – Cooperative effort and/or decision made jointly by the members of a specific group.

**Technical Manual** – A publication by test authors and publishers that provides technical and psychometric information on a test.

**Test** - Measurement instrument, procedure or device that requires scored responses from examinees; may include both selective and constructed response formats.

**Test Coordinator** – An individual charged with the oversight of the NOCTI testing program at a specific institution. This individual will need to remain informed on NOCTI's testing services, policies and procedures. NOCTI recommends that the Test Coordinator not be a faculty member, but rather an administrator or administrative staff person.

**Total Quality Management (TQM)** - TQM is a management system taken from business, which is now being applied to education. TQM applies concepts of control, quality, process, and customer services to management. The quality movement brings with it a sense of collective responsibility for learning, a habit of listening to the people we served, reliance upon data, an ethic of continuous improvement, a determination to develop fully the talent of every learner, and an acknowledgement that we are professionally accountable to one another and to those we serve.

**TURBO-DACUM** – A computer-enhanced occupational and task analysis process developed by Dr. Katherine Manley to quickly and accurately facilitate skills assessment development, curriculum design, and fundamental re-engineering of the workplace.

**Validity** - The degree to which a test or assessment measures what it is designed to measure.

**Variable** - Any quantity that can assume more than one state or numerical value.

**Weighted scoring** – A method of scoring a test in which the number of points awarded for a correct (or diagnostically relevant) response is not the same for all items in the test. In some cases, the scoring formula awards more points for one response to any item than for another.

**Web based Assessment** - Instruments (tests, surveys, questionnaires) which are deliverable over the World Wide Web and viewable through a web browser.

**Web Conference** - A form of communication, often used in test development and/or revision, involving visual participation through use of the Internet as well as audio participation through a conference call

**Work Based Learning** - Supervised student learning that takes place at a work site through unpaid or paid activities.

**Workplace Readiness Skills** - A set of personal attributes required for success at work, including the ability to identify established rules, regulations, and policies; practice cost effectiveness and time management; assume responsibility for decisions and actions; display initiative and assertiveness; and demonstrate a willingness to learn.