

Glossary of Assessment Terms

A

Ability. A characteristic that is indicative of competence in a field. (See also aptitude.)

Ability Testing. Use of standardized tests to evaluate an individual's performance in a specific area (i.e., cognitive, psychomotor, or physical functioning).

Achievement tests. Standardized tests that measure knowledge and skills in academic subject areas (i.e., math, spelling, and reading).

Accommodations. Describe changes in format, response, setting, timing, or scheduling that do not alter in any significant way what the test measures or the comparability of scores. Accommodations are designed to ensure that an assessment measures the intended construct, not the child's disability. Accommodations affect three areas of testing: 1) the administration of tests, 2) how students are allowed to respond to the items, and 3) the presentation of the tests (how the items are presented to the students on the test instrument). Accommodations may include Braille forms of a test for blind students or tests in native languages for students whose primary language is other than English.

Age Equivalent. The chronological age in a population for which a score is the median (middle) score. If children who are 10 years and 6 months old have a median score of 17 on a test, the score 17 has an age equivalent of 10-6.

Alternative assessment. Usually means an alternative to a paper and pencil test; refers to non-conventional methods of assessing achievement (e.g., work samples and portfolios).

Alternate Forms. Two or more versions of a test that are considered interchangeable, in that they measure the same constructs in the same ways, are intended for the same purposes, and are administered using the same directions.

Aptitude. An individual's ability to learn or to develop proficiency in an area if provided with appropriate education or training. Aptitude tests include tests of general academic (scholastic) ability; tests of special abilities (i.e., verbal, numerical, mechanical); tests that assess "readiness" for learning; and tests that measure ability and previous learning that are used to predict future performance.

Aptitude tests. Tests that measure an individual's collective knowledge; often used to predict learning potential. See also ability test.

Assessment. The process of testing and measuring skills and abilities. Assessments include aptitude tests, achievement tests, and screening tests.

B

Battery. A group or series of tests or subtests administered; the most common test batteries are achievement tests that include subtests in different areas.

Bell curve. See normal distribution curve.

Benchmark. Levels of academic performance used as checkpoints to monitor progress toward performance goals and/or academic standards.

C

CBM. See "Curriculum Based Measurement" below.

Ceiling. The highest level of performance or score that a test can reliably measure.

Classroom Assessment. An assessment developed, administered, and scored by a teacher to evaluate individual or classroom student performance.

Competency tests. Tests that measure proficiency in subject areas like math and English. Some states require that students pass competency tests before graduating.

Composite score. The practice of combining two or more subtest scores to create an average or composite score. For example, a reading performance score may be an average of vocabulary and reading comprehension subtest scores.

Content area. An academic subject such as math, reading, or English.

Content Standards. Expectations about what the child should know and be able to do in different subjects and grade levels; defines expected student skills and knowledge and what schools should teach.

Conversion table. A chart used to translate test scores into different measures of performance (e.g., grade equivalents and percentile ranks).

Core curriculum. Fundamental knowledge that all students are required to learn in school.

Criteria. Guidelines or rules that are used to judge performance.

Such tests usually cover relatively small units of content and are closely related to instruction. Their scores have meaning in terms of what the student knows or can do, rather than in (or in addition to) their relation to the scores made by some norm group. Frequently, the meaning is given in terms of a cutoff score, for which people who score above that point are considered to have scored adequately ("mastered" the material), while those who score below it are thought to have inadequate scores.

Criterion-Referenced Tests. The individual's performance is compared to an objective or performance standard, not to the performance of other students. Tests determine if skills have been mastered; do not compare a child's performance to that of other children.

Curriculum. Instructional plan of skills, lessons, and objectives on a particular subject; may be authored by a state, textbook publisher. A teacher typically executes this plan.

Curriculum-Based Measurement (CBM). A method to measure student progress in academic areas including math, reading, writing, and spelling. The child is tested briefly (1 to 5 minutes) each week. Scores are recorded on a graph and compared to the expected performance on the content for that year. The graph allows the teacher and parents to see quickly how the child's performance compares to

expectations.

To top

D

Derived Score. A score to which raw scores are converted by numerical transformation (e.g., conversion of raw scores to percentile ranks or standard scores).

Diagnostic Test. A test used to diagnose, analyze or identify specific areas of weakness and strength; to determine the nature of weaknesses or deficiencies; diagnostic achievement tests are used to measure skills.

E

Equivalent Forms. See alternate forms.

Expected Growth. The average change in test scores that occurs over a specific time for individuals at age or grade levels.

F

Floor. The lowest score that a test can reliably measure.

Frequency distribution. A method of displaying test scores.

G

Grade equivalents. Test scores that equate a score to a particular grade level. Example: if a child scores at the average of all fifth graders tested, the child would receive a grade equivalent score of 5.0. Use with caution.

I

Intelligence tests. Tests that measure aptitude or intellectual capacities (Examples: Wechsler Intelligence Scale for Children (WISC-III-R) and Stanford-Binet (SB:IV)).

Intelligence quotient (IQ). Score achieved on an intelligence test that identifies learning potential.

Item. A question or exercise in a test or assessment.

M

Mastery Level. The cutoff score on a criterion-referenced or mastery test; people who score at or above the cutoff score are considered to have mastered the material; mastery may be an arbitrary judgment.

Mastery Test. A test that determines whether an individual has mastered a unit of instruction or skill; a test that provides information about what an individual knows, not how his or her performance compares to the norm group.

Mean. Average score; sum of individual scores divided by the total number of scores.

Median. The middle score in a distribution or set of ranked scores; the point (score) that divides a group into two equal parts; the 50th percentile. Half the scores are below the median, and half are above it.

Mode. The score or value that occurs most often in a distribution.

Modifications. Changes in the content, format, and/or administration of a test to accommodate test takers who are unable to take the test under standard test conditions. Modifications alter what the test is designed to measure or the comparability of scores.

N

National percentile rank. Indicates the relative standing of one child when compared with others in the same grade; percentile ranks range from a low score of 1 to a high score of 99.

Normal distribution curve. A distribution of scores used to scale a test. Normal distribution curve is a bell-shaped curve with most scores in the middle and a small number of scores at the low and high ends.

Norm-referenced tests. Standardized tests designed to compare the scores of children to scores achieved by children the same age who have taken the same test. Most standardized achievement tests are norm-referenced.

[To top](#)

O

Objectives. Stated, desirable outcomes of education.

Out-of-Level Testing. Means assessing students in one grade level using versions of tests that were designed for students in other (usually lower) grade levels; may not assess the same content standards at the same levels as are assessed in the grade-level assessment.

P

Percentiles or percentile ranks (PR). Percentage of scores that fall below a point on a score distribution; for example, a score at the 75th percentile indicates that 75% of students obtained that score or lower.

Performance Standards. Definitions of what a child must do to demonstrate proficiency at specific levels in content standards.

Portfolio. A collection of work that shows progress and learning; can be designed to assess progress, learning, effort, and/or achievement.

Power Test. Measures performance unaffected by speed of response; time not critical; items usually arranged in order of increasing difficulty.

Profile. A graphic representation of an individual's scores on several tests or subtests; allows for easy identification of strengths or weaknesses across different tests or subtests.

Progress monitoring. A scientifically based practice used to assess students' academic performance and evaluate the effectiveness of instruction; can be implemented with individual students or an entire class.

R

Raw score. A raw score is the number of questions answered correctly on a test or subtest. For example, if a test has 59 items and the student gets 23 items correct, the raw score would be 23. Raw scores are converted to percentile ranks, standard scores, grade equivalent and age equivalent scores.

Reliability. The consistency with which a test measures the area being tested; describes the extent to which a test is dependable, stable, and consistent when administered to the same individuals on different occasions.

Response to Intervention (RTI). Use of research-based instruction and interventions to students who are at risk and who are suspected of having specific learning disabilities.

S

Scaled score. Scaled scores represent approximately equal units on a continuous scale; facilitate conversions to other types of scores; can use to examine change in performance over time.

Score. A specific number that results from the assessment of an individual.

Speed Test. A test in which performance is measured by the number of tasks performed in a given time. Examples are tests of typing speed and reading speed.

Standard score. Score on norm-referenced tests that are based on the bell curve and its equal distribution of scores from the average of the distribution. Standard scores are especially useful because they allow for comparison between students and comparisons of one student over time.

Standard deviation (SD). A measure of the variability of a distribution of scores. The more the scores cluster around the mean, the smaller the standard deviation. In a normal distribution, 68% of the scores fall within one standard deviation above and one standard deviation below the mean.

Standardization. A consistent set of procedures for designing, administering, and scoring an assessment. The purpose of standardization is to ensure that all individuals are assessed under the same conditions and are not influenced by different conditions.

Standardized tests. Tests that are uniformly developed, administered, and scored.

Standards. Statements that describe what students are expected to know and do in each grade and subject area; include content standards, performance standards, and benchmarks.

Stanine. A standard score between 1 to 9, with a mean of 5 and a standard deviation of 2. The first stanine is the lowest scoring group and the 9th stanine is the highest scoring group.

Subtest. A group of test items that measure a specific area (i.e., math calculation and reading comprehension). Several subtests make up a test.

[To top](#)

T

T-Score. A standard score with a mean of 50 and a standard deviation of 10. A T-score of 60 represents a score that is 1 standard deviation above the mean.

Test. A collection of questions that may be divided into subtests that measure abilities in an area or in several areas.

Test bias. The difference in test scores that is attributable to demographic variables (e.g., gender, ethnicity, and age).

V

Validity. The extent to which a test measures the skills it sets out to measure and the extent to which inferences and actions made on the basis of test scores are appropriate and accurate.

z-Score: A standard score with a mean of 0 (zero) and a standard deviation of 1.

This **Glossary of Assessment Terms** is from [Wrightslaw: From Emotions to Advocacy, 2nd Edition](#).

Sources: Center for Research on Evaluation, Standards, and Student Testing (CRESST), Graduate School of Education & Information Studies, UCLA; American Guidance Service; Harcourt, Inc.; Office of Special Education and Rehabilitative Services, U. S. Department of Education.

This page printed from: <http://www.wrightslaw.com/links/glossary.assessment.htm>