

GLOSSARY OF ASSESSMENT TERMS

Achievement Test: An objective test that measures how much a student has learned or knows about a specific subject.

Aptitude Test: An objective test that measures one or more capabilities that involve spatial reasoning, manual dexterity, clerical perception, or other capacities to learn certain behaviors. Aptitude test scores do not rely specifically on knowledge gained but may be enhanced through exposure to like activities.

Grade Equivalent: The average test score obtained by students classified at a given grade placement. For example, a grade equivalent score of 7.4 indicates that the student performed as well on the test as an average student who has been in the seventh grade for four months.

Interest Inventory: An individual survey of a person's preferences with regard to different factors related to work, such as environment, interpersonal activity, physical demand, education required, etc.

Mean: The sum of a set of scores divided by the number of scores. In lay terms, it is a synonym for the average. The value of the mean can be strongly influenced by a few extreme scores.

Median: The middle point in a set of ranked scores. It is the value that has the same number of scores above it and below it in the distribution.

Mode: The score that occurs most frequently in a distribution, no matter whether it occurs – high, low, or in the middle of the distribution.

Normal Curve Equivalent Score: A scale designed to transform percentile rankings into equal units. NCEs are suitable for computing averages, because a difference of 5 NCE units represents the same difference in achievement at any point on the NCE scale, which is not the case with percentiles. However, different NCE charts are needed for different intervals of the school year.

Percentile Rank: A type of converted score that expresses a student's score relative to the group in percentile points. Indicates the percentage of students tested who made scores equal to or lower than a given score. If a score of 82 has a percentile rank of 65, this means that 65 percent of the students who took the test had a score of 82 or lower than 82.

Reliability: The degree to which a student would obtain the same score if the test were readministered (assuming no further learning, practice effects, or other change). It is a measure of the stability or consistency of scores.

Standardization: In test construction, refers to the process of trying the test out on a group of students to determine uniform or standard scoring procedures and methods of interpretation.

Standardized Test: Contains empirically selected materials, with specific directions for administration, scoring, and interpretation. Provides data on validity and reliability, and has adequately derived norms.

Standard Score: Derived score that transforms a raw score in such a way that it has the same mean and the same standard deviation. The standard score scale is an equal-interval scale; that is, a difference of five points has the same meaning throughout the scale.

Stanine: A weighted scale divided into nine equal units that represent nine levels of performance on any particular test. The stanine is a standard score. Thus the intervals between different points on the scale are equal in terms of the number of correct test responses they represent. The mean is at stanine 5.

Validity: The extent to which a test measures what it is designed to measure. A test valid for one use may have little or no validity for another.