Standardized Test from *Dictionary of Information Science and Technology*

an assessment designed to evaluate students by comparing their work with a standard for skills and knowledge (Shaffer, 2009)

Standardized Testing

From *Encyclopedia of Human Development*

The term *standardized testing* is often used as synonymous with the kind of multiple-choice, large-scale group testing that takes place in schools and that is aimed at assessing academic achievement. This meaning of the term originated in educational settings as a way of distinguishing printed, commercially available instruments—such as the Stanford Achievement Test or the Iowa Test of Basic Skills (ITBS)—from nonstandardized, teacher-made classroom tests. However, while it is true that such testing is standardized, many other types of cognitive ability tests, as well as personality assessment instruments, can also qualify as standardized.

When educational or psychological tests or assessment instruments are described as “standardized,” that description refers to two distinct, yet interrelated, aspects of the instruments. The first aspect of standardization consists of uniformity in the administration and scoring procedures of a test. In this sense, a test is standardized if it is administered and scored according to preestablished, carefully delineated guidelines that are to be followed whenever the instrument is used—regardless of the setting in which it is used or the particular individual or group with whom it is used—in an effort to maintain fairness and objectivity. Details such as the directions to be given to examinees, the time limits for a test, the materials to be used, and the way test-takers’ questions should be addressed must be specified by the test author in the documentation that accompanies a test. The emphasis on strict control of the procedures under which the behavior samples that make up educational and psychological tests are gathered and recorded is a legacy dating back to the earliest period of experimental psychology as it developed in 19th-century Germany. At that time, experimenters became keenly aware that such things as the directions given to research participants and the conditions of the environment in which experiments took place could affect results. Thus, standardization of procedures has been a hallmark of psychological testing from its inception and constitutes an important part of the process of test development.

The second aspect of standardization in test development refers to the process wherein data based on the performance of samples of individuals are collected—under standardized conditions—and tabulated, for the purpose of setting a standard by which the performance of subsequent test takers can be evaluated. These data, in the form of score distributions, means, and standard deviations, are the norms of a test and they constitute the primary basis for norm-referenced test score interpretation; the groups from whom they are obtained are called the normative or standardization samples. In order to provide a meaningful standard of comparison, standardization samples should be representative of the kinds of test takers for whom a test is intended. Norm-referenced test score interpretation typically involves transforming the raw scores obtained by examinees into standard scores that reflect the percentile rank position of the examinee’s score within the distribution of scores of an appropriate standardization group. Thus, for example, a standard score equivalent to the
75th percentile indicates that the examinee’s level of performance on the test equaled or exceeded the performance of 75% of the individuals in the standardization sample against which the person’s score is being compared.

One of the problems inherent in norm-referenced test interpretation is that the normative frame of reference by its very nature is a relative viewpoint from which to evaluate performance and is wholly dependent on the characteristics of the standardization sample. Criterion-referenced or performance-based testing, also known as mastery testing, provides an alternative framework for test score interpretation that is becoming increasingly popular in educational settings. This framework uses preestablished standards of performance at various levels of mastery against which the performance of examinees can be evaluated in order to determine their location within the standards-based continuum of ability, knowledge, or skills that the test encompasses. In recent years, publishers of the major standardized tests used in educational settings, such as the SAT and the ITBS, have begun to incorporate both norm-referenced and criterion-referenced procedures, as well as open-ended items, into their test development in order to provide more flexibility in the way scores are used and a greater range in the behaviors they sample.

See also
SAT

Further Readings and References


Urbina, Susana

APA

Chicago

Harvard

MLA


Copyright © 2006 by Sage Publications, Inc.
APA

Chicago
https://search.credoreference.com/content/topic/standardized_tests

Harvard

MLA