

Have You Ever Wondered about the Meaning of Standardized Tests?

by the CMC Executive Board

Possibly one of the most misunderstood topics in education today is the purpose of standardized tests, such as the Stanford 9 (SAT 9), Terra Nova, California Test of Basic Skills (CTBS), or Iowa Test of Basic Skills (ITBS). While assessment serves many purposes, each particular form of testing produces different results and provides different information. What are some of the current misconceptions about standardized tests?

Misconception #1: Standardized test scores are more accurate than student report cards in measuring achievement.

Fact: Standardized tests consist of a collection of items that are in a multiple-choice format. These questions are field-tested nationwide with sample groups of students, statistically analyzed, and studied for reliability and validity. On each norm-referenced, grade-level test, students' knowledge of content is measured by test items that deal with material taught at grade level, while other items measure content that is both above and below the grade level of the student being tested. Because of the limitations of designing a test for a large number of students, not all important mathematics concepts can be measured. Moreover, only a small proportion of items actually relate to what students learn in school. Generally speaking, items can be categorized as follows:

Items that measure what a student learns in school. This involves a direct interpretation of the curriculum content that the student learns in school.

Items that measure what is learned outside of school and refer to the content of mainstream culture. Children who grow up in homes where there are opportunities to explore the world, read and develop a broad vocabulary, and experience different people, places, and things have a wealth of resources to draw on in answering these types of questions.

Items that selectively measure aspects of a child's innate intelligence, such as an affinity for using words or numbers. Children access and retain information differently and standardized tests neither match nor measure all learning styles.

A standardized test is not intended to measure mastery of content, but rather to collect results that represent a continuum of test scores ranging from 0 to 100%. To maintain the credibility of a standardized test, items that have been field-tested by a sample group and correctly answered by too many students, for example, by 80 percent or more of the students, are discarded. IF a question that focuses solely on curriculum were to be answered correctly by too many students, the end result would create a narrow range of scores. In order to ensure a broad range of scores, or spread, the test is deliberately designed so that students encounter varying levels of success in answering the different types of items.

When teachers are encouraged to focus on raising test scores, they are often tempted to devote less time to teaching curriculum and more time on test preparation. This strategy both narrows the focus of student education and the spread that the standardized test is designed to create.

Misconception #2: A student test score ranking, such as 80%, means that the student answered 80% of the test items correctly.

Fact: Standardized tests are scored on a scale, and percentile rankings reflect the comparative ranking of one student's score to a large-scale group's average performance.

The purpose of a standardized test is to generate a score spread so that comparisons can be made to "the normed group." The test is deliberately designed so that half the students will score BELOW the 50th percentile while the other half of the students score ABOVE the 50th percentile.

A nationally normed group is a sampling of students that reflects the demographics of the U.S. students in a specific age group. Considering the diverse populations of some states, an exact match to the nationally normed group cannot be guaranteed. Thus, the comparison that a standardized test can develop between the two different student population samples may foster very compelling debate concerning the kinds of educational programs that may be needed to meet the needs of varying groups of students. As educators work to refine instruction and identify program changes, they need to be aware of the possibility of differences between the two population samples—the "normed" group on which scores are based, and the "actual" students who are being tested.

So, what does a standardized test do? A well-constructed, standardized test generates a spread of scores so that comparisons to the normed group can be made. The way to guarantee that spread is to reduce the emphasis of the test content on school curriculum and to concentrate more on content that is learned outside of the classroom. Tests are deliberately designed to generate a wide range so that the students being tested can be compared to a normed group of students. Analyzing large numbers of student scores can lead to identification of patterns in the data—areas with depressed or high scores, for instance. These results have the potential to inform educational policy makers as they make funding decisions for school districts and school sites.

Misconception #3: Standardized test scores can be used to identify effective and ineffective schools.

Fact: Standardized test scores reflect the socioeconomic background of a student, more than the academic content learned in school. Given their access and exposure to the mainstream culture, students from advantaged backgrounds tend to correctly answer questions related to what's learned outside of school more often than students who come from less-advantaged situations. Students whose families have high socioeconomic levels often come from well-educated families. In their home environment, they become familiar with academic language and develop high-learning expectations that can facilitate school performance and, more specifically, school testing. Schools that exhibit high standardized test scores are not necessarily effective learning institutions. A close look at schools with high scores often reveals a consistent school population, which would predictably exhibit a high level of performance, especially if the community has a high socioeconomic profile.

Any time a single measurement tool is used to make decisions about schools, the opportunity to look carefully at the factors that comprise a good learning environment is lost. Unfortunately, standardized test scores tell us what we already know. Students who live in deficient economic circumstances need more financial resources to enrich their education. However, the current focus of rewards and punishments based on test scores, instituted by many states, pressures educators to rapidly increase scores without the long-term benefit of relevant instruction.

References

Calkins, L., K. Montgomery, and D. Santman. *A Teacher's Guide to Standardized Reading Tests: Knowledge Is Power*. Portsmouth, NH: Heinemann, 1998.

Popham, W. James. *Testing! Testing! What Every Parent Should Know About School Tests*. Boston: Allyn & Bacon, 2000

_____. "Why Standardized Tests Don't Measure Educational Quality." *Educational Leadership* (March 1999): 8-15.

Stiggins, R. and T. Knight. *But Are They Learning? A Commonsense Parents Guide to Assessment and Grading in Schools*. Portland, OR: Assessment Training Institute, 1997.