

Glossary

American College Test (ACT): A standardized examination first created in 1959 by the American College Testing Program and used primarily by schools in the Midwest and West to determine student readiness for postsecondary education. The ACT includes 215 multiple choice questions, focusing on four subject areas: English, mathematics, reading, and science. Composite scores range from 0 to 36. Approximately 40% of high school seniors in the United States take the ACT.

Advanced Placement (AP): College-level academic courses taken during high school that confer college credit if students pass standardized examinations at the end of the courses.

After-School Programs: Programs run by schools and/or community based organizations that provide recreational and learning activities for students after the end of the regular school day (usually from about 3:00 PM to 6:00 PM) or on the weekends. Some of these programs are supported through federal 21st Century Community Learning Center legislation. The legislation defines a community learning center as “an entity within a public elementary, middle or secondary school building that (1) provides educational, recreational, health, and social service programs for residents of all ages within a local community, and (2) is operated by a local educational agency (LEA) in conjunction with local governmental agencies, businesses, vocational education programs, institutions of higher education, community colleges, and cultural, recreational, and other community and human service entities.”

Alignment: Matching the skills and knowledge imparted by school or program curricula with the requirements of state standards and tests as well as the demands of postsecondary education and employment.

Applied Curriculum: A course of study that uses real-world problem solving assignments to teach theoretical concepts and academic skills.

California Achievement Test (CAT): A series of norm-referenced standardized tests used across the country to measure academic achievement of elementary, middle, and high school students in six subject areas: reading, language, spelling, mathematics, study skills, and science. Percentile scores are reported rather than raw scores.

Comparison Group: An existing collection of individuals, similar enough to the treatment group, but who do not participate in the program or initiative being studied and whose achievement is measured against the treatment group to assess the intervention’s effectiveness. For instance, a group of students in the same school but whose class does not participate in a determined program. Comparison groups are identified, but not created, by the researcher. Comparison groups are not as rigorous as Control Groups. See also Control Group, Experimental Design, Treatment Group and Matched Comparison/Control Group.

Control Group: A group of individuals who come from the same pool as the treatment group, but are assigned, preferentially through random processes (such as a lottery) not to receive the program or intervention and whose achievement is measured against that of the treatment group. Control groups are created by the researcher as part of the experimental design. See also Comparison Group, Experimental Design, Treatment Group and Matched Comparison/Control Group.

Correlation Coefficient: An index that describes the extent to which two sets of data are related or a measure of the relationship between two variables. Correlation does not imply a causal relationship. It simply indicates whether two variables (such as grades and attendance) are related.

Criterion-Referenced Test: A test that measures student achievement in relation to established skill and/or content standards rather than against the performance of other students (as in norm-referenced tests).

Effect Size: A measure of the impact of an initiative based on the difference between the mean scores of the treatment and comparison/control groups and the spread (or standard deviation) of each group's scores. See also Standard Deviation.

Experimental Design: The design of an evaluation study that randomly assigns students to treatment and control groups and holds all other factors or variables (e.g. socio-economic, demographic, environmental, etc.) constant as the students go through the education process so that the differences between the two groups can be attributed to the treatment employed (in educational research, the treatment is the program or school initiative).

High-Stakes Tests: Examinations that imply consequences for the future educational trajectory of students, teachers, principals, and schools. For students, failure carries penalties, such as not advancing to the next grade level or graduating, regardless of other measures of achievement. For schools, high failure rates might mean district take-over or revoked accreditation.

Iowa Test of Basic Skills (ITBS): A norm-referenced, standardized test used across the country to measure aptitude and achievement of students in grades K-8 in several areas: listening, word analysis, vocabulary, reading, comprehension, language skills, mathematics, social studies and science. Percentile scores are reported rather than raw scores. The Riverside Publishing Company, a subsidiary of Houghton Mifflin, publishes the ITBS.

Longitudinal Research: Evaluation that measures the effect of a program or school initiative on one group of students at different points over time. This entails tracking students' achievement while they participate in an initiative and for a number of years after they exit.

Matched Comparison/Control Group: A comparison or control group where students are similar (matched) to the treatment group in variables that are important for the research, such as race/ethnicity, age, gender, income level and academic level. The comparisons will be valid and generalizable only when the two groups (comparison or control and treatment groups) are similar or matched.

National Assessment of Educational Progress

(NAEP): Also called "the Nation's Report Card," NAEP was begun in 1969 as a continuous assessment of student knowledge and achievement in eight subject areas: reading, mathematics, science, writing, U.S. history, civics, geography, and the arts. A national sample of students in grades 4, 8 and 12 take NAEP tests. NAEP is mandated by Congress and funded through the National Center for Education Statistics in the U.S. Department of Education. Scores are reported for race/ethnicity and other subgroups of students nationally and by state but not for individual students or schools.

Normal Curve Equivalent (NCE): A standardized scale of scores developed by the U.S. Department of Education that allows comparison between different types of tests and different groups of students taking the same test. NCE scores have a normal curve distribution with a mean of 50 and standard deviation of 21.

Norm-Referenced Test: A test that measures the performance of individuals against the mean performance of the other students taking the test rather than against a set of skill or content standards. Scores on norm-referenced tests are relative and usually reported as percentiles. (See also Criterion-Referenced tests.)

Percentile: The standing of an individual in relation to a larger group of students taking the same test (e.g., a student scoring at the 75th percentile scored higher than 75% of the students taking the test, but did not necessarily get $\frac{3}{4}$ of the answers on the test correct). [Percentiles are values that divide a sample of data into one hundred groups containing equal numbers of observations. For example, 50% of the data values lie below the 50th percentile.]

Sample Size: The number of students included in an experiment or evaluation, usually smaller than the total number of students participating in the program or school initiative but large enough to represent the entire group.

Scholastic Aptitude Test (SAT): A standardized examination first developed by the College Board in 1926 to determine student readiness for postsecondary education. Today, the SAT focuses on two subject areas, measuring verbal and

mathematical ability. It is one of the most popular tests used by college admissions officials (in conjunction with other application materials) to determine whether to accept prospective students. Scores range from 200 points to 800 points on the Verbal and Mathematics sections of the test with a maximum potential score of 1600. Approximately 40% of high school seniors in the United States take the SAT.

Small Learning Communities: The organization of a school or youth program that is small enough to allow for personalized attention by staff and teachers for each student. The U.S. Department of Education supports small learning communities through Part A of Title X of the Elementary and Secondary Education Act (ESEA), which provides grants to implement career academies, schools-within-schools, mentoring, career clusters and other strategies to restructure large schools.

Standard Deviation: Standard deviation is a measure of the spread or dispersion of a set of data. The more widely the values are spread out, the larger the standard deviation. For example, if a group of students was given two exams and their scores varied from 30 to 98 on the first exam and from 80 to 98 on the second exam, the standard deviation is larger for the first exam.

Stanford-9 (SAT-9): A norm-referenced, standardized test used in schools nationwide to measure student achievement in grades K-12. For elementary and middle school students, the test focuses on vocabulary, reading, writing, spelling, and math. The subject areas for high school students are reading, writing, history/social science, math, and science. The Harcourt Brace Educational Measurement division publishes the SAT-9. Percentile scores are reported rather than raw scores.

Texas Assessment of Academic Skills (TAAS): A criterion-reference test administered annually by the Texas Education Agency that measures student achievement in reading and math (for grades 3-8 and 12), in writing (for grades 4, 8, and 12), and in science and social studies (grade 8). Spanish language versions are available for grades 3-6. Texas students must pass the TAAS to graduate with a high school diploma.

Treatment Group: In education research, this is the collection of students who participate in a program or school initiative. Their attitudes or achievement outcomes are often measured against control or comparison groups to determine the effectiveness of the program or school initiative. See also Control Group and Comparison Group.

Note: For more information on testing and educational research terms, see Gerald W. Bracey, "Thinking About Tests and Testing: A Short Primer in 'Assessment Literacy'" (Washington: American Youth Policy Forum, 2000). Available on line at www.aypf.org/BraceyRep.pdf.