

Association of American Publishers

students

enhanced learning

information

# STANDARDIZED ASSESSMENT:

A PRIMER (Revised Edition)

successful schools

high standards





## Introduction

Educational testing is an essential activity in every school, in every school district and in every state. Standardized tests are used to evaluate students and schools; to help improve teaching and learning; and to generate important data from which policy decisions can be made. Standardized test results often are a major force in shaping public perceptions about the performance of our students and the quality of our schools. With the recent passage of the No Child Left Behind Act of 2001 (NCLB Act),<sup>1</sup> which requires tests to be the primary measure of school accountability, testing has taken on added significance.

A number of national surveys show testing has strong support among the public. For example, a survey by the Association of American Publishers found overwhelming parental support for standardized testing.<sup>2</sup> According to

the survey, a large majority of parents (83 percent) believe that standardized tests provide important information about their children's educational progress, and nine out of 10 parents said they want comparative data about their children and the schools they attend. More than two thirds of all parents (67 percent) surveyed said they would like to receive standardized test results for their children in every grade.

Other surveys also demonstrate support for educational testing, echoing AAP's findings. Public Agenda's "Reality Check" opinion survey for 2002 found that "Even as students nationwide face more testing and higher hurdles for promotion and graduation, very few seem apprehensive about school or unnerved by what is currently being asked of them" and that there is "Broad agreement...that testing has genuine

benefits."<sup>3</sup> A survey released by the Business Roundtable in September, 2000, found that 85 percent of the American public says statewide tests



are useful to schools in evaluating how well students are performing.<sup>4</sup> Eighty-three percent of respondents said test scores are very useful for parents and the community in evaluating how well schools perform, and 78 percent of the public said test

scores help parents keep up with how their own children are doing in school. Finally, a survey conducted by the American Association of School Administrators in 1999 found that "79 percent of parents surveyed agreed that standardized tests are necessary to measure what students have learned."<sup>5</sup>

In order to provide policy makers and the public with information about the role of testing, what tests can do, and how testing systems should be designed, created, and administered, the Association of American Publishers in 2000 created *STANDARDIZED ASSESSMENT: A PRIMER*. This revised version of the brochure provides answers to commonly asked questions about tests and testing at the elementary and secondary school levels, and reviews the key elements of the NCLB Act that are essential to effective testing programs.

<sup>1</sup> Public Law No. 107-110.

<sup>2</sup> AAP, July 2000. The national survey of 1,023 parents of school-age children was conducted by JD Franz Research, Inc. of Sacramento, CA in April, May, and June 2000. The survey had a margin of error of 3.5 percent. The entire survey can be viewed on AAP's web site at [www.publishers.org](http://www.publishers.org).

<sup>3</sup> Public Agenda, 2002. <http://www.publicagenda.org/specials/rcheck2002/reality.htm>

<sup>4</sup> Business Roundtable, Sept. 2000.

<sup>5</sup> *Leadership News*, July 2000, American Association of School Administrators.

## Why Test?

The purpose of tests is to provide educators, students, parents, and policy makers with information that is valid, fair, and reliable.

Standardized tests provide information that helps support four critically important tasks for educators and the public:

1. Identify the instructional needs of individual students so educators can respond with effective, targeted teaching and appropriate instructional materials;
2. Judge students' proficiency in essential basic skills and challenging standards and measure their educational growth over time;
3. Evaluate the effectiveness of educational programs; and,
4. Monitor schools for educational accountability including under the NCLB Act.

In sum, tests provide information to help students learn more successfully, teachers teach more effectively, and schools to be more accountable.

By far the majority of respondents (84 percent) said standardized test scores measure all, most, or some of the things that are important to them about their children's education. Two-thirds of respondents (67 percent) said they would like to receive standardized test results for their children in every grade.

*AAP Survey of Parents*

There are limits to testing, however. Tests are a necessary but not the exclusive means to evaluate current achievement and students' growth in skills. What may be tested is not, and cannot be, inclusive of all of the desired outcomes of instruction. Tests should be considered a means to an end and not ends in themselves. Tests should be used in combination with other important types of information such as teacher judgments of student work and classroom performance plus other individual and group assessments, to measure achievement and growth.

## What Is A "Standardized Test"?

A *standardized achievement test* is, simply, a test that is developed using standard procedures and is then administered and scored in a consistent manner for all test takers. Students respond to identical or very similar questions under the same conditions and test directions. The standardization of test questions, directions, conditions of testing, and scoring is needed to make test scores comparable and to assure, as much as possible, that test takers have equal, unbiased opportunities to demonstrate what they know and can do.


Standardization can apply to any type or format of test. However, some types of educational tests such as classroom and teacher-developed tests are not usually considered to be "standardized" tests because they are given under varying conditions and are scored using variable rules.

Standardized tests may be used for a variety of purposes. One purpose of testing is to enable educators to make high-stakes decisions about individual students through measures such as high school graduation tests. In contrast, the annual testing provisions of the NCLB Act are used to inform schools, teachers, and parents about student improvement in the classroom and to hold schools and states accountable for such improvement.

Parents who responded to the AAP survey rated standardized tests as trustworthy, compared to other measures of their child's learning in school:

*Fairness*—63 percent responded that standardized tests are as fair or more fair than other measures  
*Consistency*—57 percent responded that standardized tests are as consistent or more consistent than other measures  
*Unbiased*—58 percent responded that standardized tests are as unbiased or more unbiased as other measures





## What Types of Information Can Tests Provide?

Tests can provide information on individual student or group performance that can be interpreted and used in many different ways. There are currently educational testing systems at the national, state, and local levels. The National Assessment of Educational Progress (NAEP) is a federal testing program that generates periodic “snapshots” nationally and within states of how small samples of students perform in a given subject at a particular grade level. States and districts use


standardized tests, usually administered to all children rather than by samples, to generate information that teachers, parents, and policy makers need to make decisions about schools and students. Most standardized testing programs at the school and classroom levels are designed to help teachers improve their teaching of specific subjects and skills, thus improving individual student learning.



## What Formats of Questions May Be On A Standardized Test?

Standardized tests can have a variety of questions and formats. Any format can present challenging questions that test higher order skills, as well as fundamental skills. A standardized test may include more than one format of question. The most common formats are:

**Multiple-choice questions:** Many standardized tests require students to select a single correct response to each test question (called “items”) from among a small number of specific choices. This format—called “multiple choice” or “selected response”—is efficient, practical, and usually produces highly reliable results. Multiple-choice tests offer the advantages of objectivity and uniformity in scoring, ease of administration, and low cost.



## What Is The Difference Between A Criterion-Referenced Test and A Norm-Referenced Test?

**Performance assessment questions:** Performance assessments require students to generate a response to a question rather than choosing from a set of responses provided to them. Examples include exhibitions, investigations, demonstrations, written or oral responses, journals, and portfolios. Performance assessments can be given and scored according to standard procedures and rules so that a test containing performance assessment questions is a standardized test. Performance assessments typically focus on the process of problem solving rather than on answers or solutions. Tests including performance assessments, however, are generally less reliable, more difficult to score, and more costly than tests using multiple-choice items.

**Constructed-response questions:** Constructed-response items may be one type of performance assessment, in which students are given the opportunity to fill-in-a-blank or provide a brief written response to a question, rather than select from an array of possible answers. Constructed-response questions are often included, along with multiple-choice questions, on a test to obtain additional and different types of information about what a student knows or can do.

All standardized tests now administered to elementary and secondary school students measure student achievement against a set of academic standards or curricular objectives. The standards may be common among the states and major national academic organizations, thus enabling national comparisons. Or, the standards may be local standards chosen by the school district or state, which may only allow local comparisons among students in a district or state.

There are many ways to report and interpret the results of a standardized test. One way is based on specific criteria, such as academic skills or objectives and academic achievement standards developed at the state or local level. For example, “She has demonstrated mastery of reading at the third-grade level” is a determination made by a criterion-referenced test (CRT). A standardized test also can describe a student’s performance compared to other students nationally or locally. For example, “He reads better than 90 percent of fourth grade students nationally” is a determina-

tion made from a norm-referenced test (NRT). A student's score on a CRT using local academic standards is intended to be compared only with other students who have taken the same test. In contrast, a student's scores on an NRT can show performance on academic standards and also enable comparisons with students both locally and nationally. When a local CRT is used with a national NRT, the results can be interpreted together to obtain more comprehensive information about a student's performance. For example, "She is 'proficient' on a state-mandated CRT and is performing at an academic level that is better than 70 percent of students nationwide."



## How Are Standardized Tests Used?

Seventy-four percent of parents responding to the AAP survey said they get information about their children's progress in school from scores on standardized tests. When asked how they find out information about the school their children attend, the most common response (56 percent) was "standardized tests."

Information from standardized tests can be used for many purposes. These purposes may include:

### **Supporting instructional decisions for individual students by identifying their instructional needs.**

A test may be used to diagnose a student's strengths and weaknesses, thus allowing the teacher or school to choose effective instructional programs for the student.

**Demonstrating students' proficiency in basic skills and their ability to meet academic standards.** Test results are used by states to demonstrate individual student mastery of specified levels of achievement.



**Informing parents and the public about school and student performance.** States administer standardized assessments and report the results, in part to inform the public about how well the schools and their students are progressing over time and compared to other localities or schools. Many states and districts publish annual report cards on school districts and individual schools. The results of the tests can motivate education reform by informing and influencing parents to take action to improve the quality of local schools.

### **Holding schools and educators accountable for student performance on tests aligned to high standards of what students should know and be able to do.**

Consequences are often attached to test results and may include school improvement plans, technical assistance, increased or decreased funding for schools, salary bonuses, promotions, loss of accreditation and takeovers of local schools by the

state. Such consequences are used to leverage change at the school and classroom level.

**Evaluating programs.** Many federal and state education programs use standardized tests to determine if public policy objectives are being achieved, and if public funds are well-spent.

**Determining rewards and sanctions.** Tests may be used for high-stakes purposes with rewards and sanctions to make decisions about individual students, such as placement in specific programs or classes, graduation from high school, or promotion to the next grade.



Close to two-thirds of parents (63 percent) told AAP there are benefits from the information provided by standardized tests. Furthermore, the great majority—75 percent—of respondents wanted answers to two questions: (1) *How well is my child doing and is he/she making progress?*, and (2) *What are the areas in which my child needs help, so I can help him/her?*

No single test can accomplish all of the goals and objectives of diverse educational systems. It is important both legally and technically not to put all of the weight on a single test when making important decisions about students and schools. Rather, there must be multiple measures or indicators of performance to support important decisions.



## Educational Assessment and The No Child Left Behind Act of 2001

For many decades state legislatures, state education agencies and local school districts have set policies regarding educational assessment. In 1965, the federal government became involved in providing funds to help districts serve the needs of disadvantaged students. Originally, Title I of the Elementary and Secondary Education Act (ESEA) required all children in the program to be tested annually in every grade to determine if they made progress during the school year toward achieving grade level academic performance. Gains in performance were expected and remedial action was required if gains were not achieved.

The Improving America's Schools Act of 1994 (IASA) amended the ESEA. The legislation changed the focus from an emphasis on individual children in the Title I program to aggregate performance of all children measured against state academic content and achievement standards in reading/language arts and mathematics, using tests aligned to those standards. Tests had to be administered at least once in each of three grade spans: 3-5, 6-9, and 10-12.

## Key NCLB Act Provisions

In late 2001, Congress reauthorized the ESEA as the NCLB Act. The new law brought testing to an entirely new level of importance by requiring that tests be the primary means for determining school accountability for a new requirement that all students perform at a state's "proficient" level by 2013. It expanded the prior testing requirements to include annual testing in reading/language arts and mathematics, in each of grades 3-8 and once at the high school level. Beginning in 2007, the NCLB Act will also require science assessments at least once in each of three grade spans: 3-5, 6-9, and 10-12. It requires more extensive corrective actions for schools whose students failed to make annual improvement—known as "adequate yearly progress"—on the state's test. The NCLB Act created significant changes and challenges for American students, teachers, and parents.

Overall, the purpose of the new law was to raise expectations and to help educationally disadvantaged children achieve to the same high state academic achievement standards as all other students. The NCLB Act is the most significant attempt to date by the federal government to hold states and local school districts accountable for educating all of their students to high standards and to assure the effective use of federal Title I funds.

In requiring states to administer annual reading, math, and science assessments, the NCLB Act directs states—not Congress or the U.S. Department of Education—to select and design tests of their choosing, aligned to their own academic content and achievement standards. States must set at least three levels of passing scores for what students are supposed to know and be able to do. The law refers to these as "achievement levels" and calls them "basic," "proficient," and "advanced" but states may give them other names. States have until the 2005–2006 school year to expand and implement their math and reading assessments in all required grades. States must develop science standards by the 2005-2006 school year and then implement science assessments by the 2007–2008 school year in one grade in each of the three grade spans: 3–5, 6–9, and 10–12.

In 2002, the Department of Education issued regulations implementing the standards and assessment provisions of the NCLB Act that provide states significant flexibility in designing their assessment systems. The regulations allow state assessment systems to use a criterion-referenced test, a norm-referenced test, or a combination of the

two types to meet the testing requirements of the NCLB Act. If a state chooses to use only an NRT, it must be augmented as necessary with additional items to measure accurately the state standards. A state may also use “combined assessment systems” that include both state and local assessments, if it can demonstrate that the system has a “rational and coherent design.” The NCLB Act also requires all states to participate in the National Assessment of Educational Progress (NAEP). State NAEP tests will be administered in reading and math every other year at the fourth and eighth grade levels to only a sample of students in each state. In requiring states and school districts to make “adequate yearly progress” (AYP), the Act mandates that each state’s definition of “adequate yearly progress” apply both to disadvantaged students as well as to the overall student population. By setting an identical, high goal for all students, the purpose of the AYP provision is to close the achievement gap between disadvantaged children and other students.

The NCLB Act requires that all students achieve at the state’s “proficient” level on state reading and math academic assessments by the 2013-2014 school year, with the state setting annual improvement targets to assure that the goal is reached. Annual targets also must be set for specific subgroups of students to assure that overall

improvements do not mask poor achievement by economically disadvantaged students or racial or ethnic minorities. Schools and school districts that fail to meet state-defined levels of adequate yearly progress must take a variety of corrective actions, depending on how long they have been in need of improvement.

The law also requires states and school districts to report student performance for all students and *disaggregated* or separately reported performance by specific subgroups in annual, public report cards beginning with the 2002-2003 school year. The report cards must contain information about the number and percent of students achieving proficiency on state standards; graduation rates; number and names of schools identified for improvement; the professional qualifications of teachers; comparison of students at basic, proficient, and advanced levels of academic achievement on state assessments; and comparison of the actual academic achievement levels for all groups of students compared to the annual objectives for such groups. The disaggregated reporting categories must include race, ethnicity, gender, disability status, migrant status, English proficiency status, and disadvantaged economic status.

While the NCLB Act sets new federal requirements regarding state and local assessment activities, it does not pre-empt the important role

states and school districts play in educational assessment. In fact, the law places more emphasis on state and local assessment activity than ever before—especially on assessments that will assist teachers in ensuring a high level of learning during the school year. The close working relationships that have grown over the years between test publishers, states, and school districts will take on even more significance in the NCLB Act era.



## How Are Effective Testing Programs Designed?

To effectively design a standardized testing program, including those required under the NCLB Act, a plan that answers the following questions is the first and most important step.

**What is the purpose of the new assessment program?** Will it be used to help teachers better focus instruction for individual students, measure student progress, evaluate programs, or determine accountability? What type or types of information will the assessment generate and how will the information be used? Will it be used for high stakes purposes? If so, does it meet legal standards and accepted professional and technical standards?

**On what education goals and standards will the assessment program be based?** Under the NCLB Act, states must adopt academic and content standards that specify what children should know and be able to do. Local school districts may also establish their own standards. Next, curricula need to be adopted and instructional materials selected to help teachers assist their students in meeting the standards. Finally, tests must be developed to measure students’ progress toward meeting the estab-

lished academic standards. In sum, assessment plays an integral role in reforming our schools.

**Have steps been taken to ensure that the new assessment is valid, fair and reliable?** Assessments should be designed, pilot tested, and used following nationally accepted technical standards such as those developed by the American Psychological Association, the American Educational Research Association and the National Council on Measurement in Education.<sup>6</sup> The NCLB Act requires that assessments must be consistent with these standards, and that there be full participation by all students, including accommodations for students with disabilities and limited English proficient students. Parents should check with their local districts to see what accommodations are available.

**Will the new assessment program recognize the important role of multiple measures?** Just as different tests provide different information, no single test can tell us all we need to know about a student. The use of multiple measures as the basis for high-stakes decisions is one of the most important prerequisites for ensuring the fairness of new assessment programs.

**How will the new testing program be introduced and communicated to teachers, parents and students?**

Public understanding and trust must be created and maintained in order to ensure the long-term viability of any assessment system. This can be accomplished by clear communication that starts early in the process and is maintained on a continuous basis.

Eight in 10 respondents told AAP they would like to receive a simple but comprehensive explanation of how to interpret standardized test scores.



## A Final Word

Educational testing offers enormous benefits for parents, students, educators, employers, and policy makers. However, those benefits can only be realized by using tests of the highest technical quality and for the purposes for which they were designed. Today, the nation's test publishers work in close cooperation with states and school districts to provide the technical expertise and professional assistance necessary to develop, score, and interpret high-quality assessments. Nationwide, these close partnerships have led to new assessment systems that are meeting and surpassing the needs of teachers, parents, students, policy makers, and federal law.



## Where Can I Go For More Information?

The Association of American Publishers School Division is the principal trade association of the educational publishing industry, which includes all the major test publishing companies. AAP and its member publishers assist policy makers and educators in every state on a variety of issues related to educational publishing. For more information, please contact:

Stephen Driesler  
Executive Director  
School Division  
Association of American Publishers  
50 F St., N.W., Suite 400  
Washington, DC 20001  
202/220-4549



## AAP School Division Test Committee

**CTB/McGraw-Hill,**  
[www.ctb.com](http://www.ctb.com)

**Educational Testing Service,**  
[www.ets.org](http://www.ets.org)

**Harcourt Educational Measurement,** [www.hemweb.com](http://www.hemweb.com)

**NCS/Pearson,**  
[www.ncspearson.com](http://www.ncspearson.com)

**Publishers Resource Group, Inc.,**  
[www.prgaustin.com](http://www.prgaustin.com)

**Riverside Publishing Company,**  
[www.riverpub.com](http://www.riverpub.com)

<sup>6</sup>American Psychological Association, the American Educational Research Association and the National Council on Measurement in Education, *Standards for Educational and Psychological Testing* (1999)



Association of American Publishers  
50 F St., N.W., Suite 400  
Washington, DC 20001  
202/347-3375  
[www.publishers.org](http://www.publishers.org)