
HIGH-STAKES TESTING AND NO CHILD LEFT BEHIND: INFORMATION AND STRATEGIES FOR EDUCATORS



NATIONAL
ASSOCIATION OF
SCHOOL
PSYCHOLOGISTS

By Jeffery P. Braden, PhD, North Carolina State University; &
Jennifer L. Schroeder, PhD, NCSP, Dallas (TX) Independent School District

The No Child Left Behind Act of 2001 (NCLB) is the largest federal funding program for education in U.S. history. NCLB is the most recent reauthorization of the Elementary and Secondary Education Act (ESEA), which originated in 1965 as part of the War on Poverty. The primary focus of ESEA/NCLB is Title I, which allocates federal funds to help states educate economically disadvantaged children.

Schools in which 40% or more of the student body are below the poverty line are eligible for Title I funds. Unlike special education legislation, Title I programs do not establish eligibility criteria for students; eligible schools may serve any (or all) students at the site using Title I funds. NCLB has eight other titles in addition to Title I, which address topics such as educator quality, language instruction for Limited English Proficient (LEP) and immigrant students, school safety, innovation, assessments, and American Indian education.

Title I is the most important component of NCLB. This is true for two reasons: (a) The vast majority of funds are committed to Title I and (b) Title I requires substantial state accountability for improved student learning as reflected on statewide testing.

Therefore, this handout is devoted to describing NCLB Title I legislation related to assessment and accountability, the consequences or stakes attached to assessment, the intended and unintended consequences of large-scale assessment, and recommended responses.

NCLB Assessment and Accountability

NCLB expands previous assessment requirements, and substantially increases both the expectations and consequences for accountability. Previous versions of ESEA required states to develop educational standards that identified what states expected students to know and do. States were also required to test students in an elementary, middle or junior high school, and high school grade annually (e.g., grades 4, 8, and 10), and to require schools to show improvement on assessments known as Adequate Yearly Progress (AYP). States could determine the assessments, the subject matter assessed, AYP expectations, and what to do about schools that do not meet AYP expectations.

NCLB is more expansive and explicit than previous legislation in defining assessment and accountability. The most important changes include the following:

- Annual reading and mathematics testing is required in grades 3–8, and one high school grade, by the 2005–2006 school year. Annual science assessment in an elementary, middle school or junior high school, and high school grade is added in 2007–2008.
- AYP is defined as *progress* toward meeting the goal of 100% of all children in a state to meet state proficiency standards by 2014. Other indicators (e.g., attendance) may also be used to track progress, but achievement is considered the essential goal.
- Schools are required to track and meet AYP for identifiable subgroups, including groups defined by race/ethnicity, poverty, gender, disability, and English proficiency. States must also ensure that they include at least 95% of the students in each category in annual assessments.
- Federal funds will be withheld from states failing to meet progress and inclusion requirements; states must provide funding and “corrective action” to schools failing to meet AYP for two consecutive years (i.e., identified as “In Need of Improvement”). Schools needing improvement for more than two consecutive years must consider restructuring in major ways (e.g., reconstitution as a charter school).

- States are required to inform parents of the AYP status of every school, and provide parents in schools needing improvement with the right to transfer their children to a school within the district that meets AYP with the district providing free transportation. Districts must also provide and supplemental services such as private tutoring for economically disadvantaged students attending schools in need of improvement for more than one year.
- Schools in need of improvement must develop plans to improve. The plans must incorporate instructional strategies from “scientifically based research.” These schools also must spend at least 10% of NCLB funds on professional development.

Defining Adequate Yearly Progress

NCLB requires states to identify schools not making AYP, and specifies a series of consequences for schools failing to meet AYP for two or more consecutive years. However, NCLB allows states to define AYP for schools, provided that the AYP goals consider the school’s level of performance in 2002, and the goal of having all (100%) of the school’s students proficient by 2014. There are three approaches to defining AYP:

- *Fixed or absolute methods:* In which the proportion of proficient students must be equal to or greater than a fixed standard (e.g., 70% in 2002, 85% in 2008, 100% in 2014).
- *Cross-sectional methods:* In which scores from one or more years (e.g., the average of the current plus past year) are compared to scores from other years (e.g., the average of scores from the prior two years) to see if the school is making progress toward proficiency. Often, this is accompanied by the expectation that schools starting with lower scores will need to make more rapid progress than schools with higher scores.
- *Longitudinal or value added methods:* In which changes in scores for the same students from one year to the next help decide whether the school is making progress.

Pros and cons. There are significant pros and cons for each approach. The fixed standards are perhaps the easiest to understand because schools either meet or do not meet the standard. However, fixed standards are not sensitive to differences between schools, changes in student characteristics, and are most susceptible to error because of small samples.

Cross-sectional methods are sensitive to between-school differences, because they adjust for the school’s

initial starting point. Cross-sectional methods are also less susceptible to sampling error, because they can combine results over multiple years. However, cross-sectional methods will vary if student characteristics change from year to year.

Longitudinal methods are most likely to reflect school quality because they adjust for differences in students. However, longitudinal methods do not reflect absolute levels of proficiency as demanded by NCLB. Moreover, longitudinal methods may produce similar results to other methods (e.g., gains are generally smaller for low-performing students than for high-performing students, so schools with many low-performing students are likely to produce smaller average gains than schools with high-performing students).

Most states combine one or more of these methods to determine AYP. For example, a state may set an absolute standard (schools must have at least 50% of students proficient) and also require improvement from previous year data (i.e., a cross sectional method) and a minimum average gain from grade to grade (i.e., value-added or longitudinal method). This means a school might fail NCLB standards (because it did not meet the absolute standard for proficiency in a year), but could be identified as exemplary by the state based on its cross-sectional or longitudinal improvements. The opposite can also happen: A school may exceed NCLB standards, but fail state standards because it did not show gains from previous years (cross-sectional improvement) or from year-to-year in the same school (longitudinal improvement).

Challenges of AYP. There are some fundamental challenges in addressing AYP:

- AYP goals must be *sufficiently high* to induce progress toward 100% proficiency by 2014.
- AYP methods must be *responsive to improvement* by rewarding gains even among low-performing schools.
- AYP criteria should be *resistant to corruption*. Corruption refers to activities that alter the meaning of test scores so that they do not accurately reflect student knowledge. Corruption can be intentional (e.g., cheating) or unintentional (e.g., practicing test-taking skills).
- AYP should be *sensitive and responsive to differences* among students and schools. It is difficult to reconcile these competing goals. For example, Colorado allows students with disabilities to be classified as proficient even if they are performing at a level lower than their non-disabled peers. Consequently, Colorado’s AYP criteria will allow even low-performing students with disabilities

to contribute to AYP goals, but these same criteria may encourage corruption by pressuring educators to identify low-performing students as having disabilities so that they can count in favor of AYP goals.

High Stakes, But for Whom

Educators and others often describe state assessments as *high-stakes testing*, but this description confuses the stakes with the stakeholders. Briefly, high-stakes testing means that significant (high) consequences (stakes) are influenced by test scores.

Consequences of failure to meet AYP. NCLB accountability requirements provide severe economic consequences for states failing to meet AYP and assessment inclusion goals. Consequently, states develop consequences for districts, schools, and their employees. Do not assume that NCLB requires states to withdraw funding from troubled schools. In fact, NCLB requires states to give money to schools not meeting AYP and inclusion targets, and allows states to provide monetary rewards to schools exceeding expectations.

A more likely consequence for failure to meet AYP goals is a loss of autonomy. Educators at a district or school not making AYP are likely to lose the opportunity to determine teaching materials and methods, governance structures, and other forms of professional autonomy, such as the privacy and freedoms common to most individual classroom teachers. NCLB requires educators at schools to restructure their practices, and to adopt practices supported by scientifically based research, which implies substantial changes in organization and behavior.

What stakes are not in NCLB. It is important to note what stakes are *not* in NCLB. Nothing in NCLB requires states to use tests for:

- Student promotion, retention, or graduation.
- Teacher contract renewal.
- Linking teacher pay to test performance.

States or districts may opt to invoke such consequences for students or educators, but NCLB does not require such stakes. In 2002, at least 19 states required graduation tests, and at least 6 used tests for student promotion. However, the high stakes in NCLB are mostly targeted educational agencies and, by extension, educators' professional practices. The only direct consequences for other educational stakeholders are for parents and students at schools failing to meet AYP goals for two or more consecutive years (i.e., intra-district transfer and supplementary educational services).

Intended and Unintended Consequences of High-Stakes Testing

Large-scale assessments create consequences for school systems and individuals. Some consequences are intended, and others are not. However, ethics require educators and others to consider all consequences of assessment and not just intended ones.

Intended consequences. Although the main goal underlying the use of large-scale assessments is to increase student achievement, there are other consequences that proponents of large-scale assessments often have in mind when assessments are implemented. These include:

- Improving instructional focus on things stakeholders feel students should learn (i.e., aligning instruction to standards).
- Increasing instructional efficiency and resource allocation to improve student performance.
- Increasing motivation for students, teachers and parents to perform, teach, and support.
- Reducing the achievement gap between majority and historically underserved or low-performing groups.
- Increasing the use of evidence-based instructional methods (and decreasing ineffective approaches).

Unintended consequences. In addition to the intended consequences of assessments, unintended consequences affecting schools and individuals may include:

- Narrow curricular focus to teach only what is tested (e.g., eliminating music classes).
- Academic demoralization among teachers, students, and parents following repeated test failure.
- Anxiety among students, teachers, and parents.
- Corruption, such as cheating or teaching to the test (rather than to the standards).
- Inappropriate resource allocation, such as targeting students near the cutoff while ignoring students not near cutoff levels.
- Exclusive use of a single data sources (i.e., test) to make high-stakes decisions.

Outcomes. The evidence is mixed with respect to the degree to which intended and unintended consequences occur. Whereas some research suggests states with high stakes (i.e., significant consequences for students and schools) tests do no better on National Assessments of Educational Progress (NAEP), other analyses of NAEP data show higher stakes are associated with decreasing minority achievement gaps

within states. Ultimately, individual educators will influence the degree to which intended versus unintended consequences follow from testing. The next section describes ways in which educators can increase the beneficial consequences, and minimize the harmful consequences, of high-stakes testing

Promoting Constructive Large-Scale Assessments

Educators can take steps to ensure high-stakes testing programs do more good than harm. Although there are certainly roles for individuals and organizations to influence policies, the focus should be on individual educators and their schools. However, educational stakeholders may be uninformed, or misinformed, about testing programs, results, and consequences. When educators lack this knowledge, there may be a greater impact of the unintended consequences of the tests.

The following suggestions may help educators and policymakers to make better decisions about how to respond to high-stakes testing and NCLB accountability provisions:

Promote assessment literacy. Some educational stakeholders may believe that most children are at or above grade level, or that teacher judgments are more reliable and less biased than standardized tests. Helping recipients (e.g., students, parents, teachers) to accurately interpret the results of state tests increases good decision making and decreases harmful responses to testing.

Use results to get results. Although test scores often arrive too late in the year to help plan instruction for individual students, scores can provide valuable insights for strategic instructional planning. Noting which instructional objectives are more or less likely to be mastered in each grade can lead educators to analyze and reflect on their educational practices for next year's class. Such annual planning can include strategic decisions related to instructional practices, materials, time on task, sequencing instruction within and across grades, and other decisions to increase student learning. Professionals who understand how to provide and interpret these data can help teachers use data to transform and improve student learning.

Feature the facts. Many people misunderstand educational accountability systems and tests. For example, some educators believe that excluding students with disabilities (or other historically low-performing students) will improve school performance, but NCLB requires test participation for all students, and provides specific consequences for exclusion of students from testing. Another misconception is that

testing is extremely time consuming; in fact, most state test programs consume less than 2% of the instructional time in any given grade/year.

However, facts may also be critical of assessments. For example, accountability systems often favor schools with wealthy/high performing students. A balanced, accurate presentation of facts is more likely to improve decision-making.

Bring science to schools. NCLB explicitly requires schools to use methods supported by scientifically based research, in part because such research suggests that there is relatively widespread use of inefficient, unsupported instructional practices. Bringing scientific rigor (e.g., knowledge of measurement, experimental design, literature synthesis) to school improvement plans can help schools adopt high quality strategies (e.g., explicit phonemic awareness instruction) to improve student learning, and avoid low quality strategies (e.g., test preparation programs) to change student test scores.

Help people cope. Tests induce stress by changing routines, creating uncertainty, and, for some people in some settings, changing lives through promotion, graduation, or, in some cases, radical changes to failing schools (e.g., closure, reconstitution). Educators need to have professional services available to help them understand the testing process, identify and implement evidence-based methods of instruction, and receive emotional support to help them cope with the stresses of testing.

Resources

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). *Standards for educational and psychological testing* (3rd ed.). Washington, DC: Author. ISBN: 0935302255.
- Amrein, A. L., & Berliner, D. C. (2002). *An analysis of some unintended and negative consequences of high-stakes testing*. Tempe, AZ: Arizona State University, Educational Policy Studies Laboratory.
- Bracey, G. (2000). *High-stakes testing*. Tempe, AZ: Arizona State University, Educational Policy Studies Laboratory. Available: www.asu.edu/educ/eps/EPRU/documents/cerai-00-32.htm
- Braden, J. P. (2002). Educational accountability: High-stakes testing and educational reform. In A. Thomas & J. Grimes. (Eds.), *Best practices in school psychology IV* (pp. 301–319). Bethesda, MD: National Association of School Psychologists. ISBN:0-932955-85-1.

- Gratz, D. B. (2000). High standards for whom? *Phi Delta Kappa*, 81, pp. 681–687. Available: www.pdkintl.org/kappan/kgra0005.htm
- National Association of School Psychologists. (2003). Position statement: Using large-scale assessments for high-stakes decisions. Available: www.nasponline.org/information/pospaper_largescale.html
- Thum, Y. M. (2003). *No Child Left Behind: Methodological challenges and recommendations for measuring adequate yearly progress*. Technical Report 590, National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA. Available: www.cresst.org/products/reports_set.htm

Websites

- National Center for Educational Outcomes, University of Minnesota—www.education.umn.edu/NCEO
- National Center for Research on Evaluation, Standards, and Student Testing—www.cresst.org
- National Reading Panel—www.nationalreadingpanel.org
- No Child Left Behind home page—www.nochildleftbehind.gov
- Thomas B. Fordham Foundation (resources related to standards and assessment)—www.edexcellence.net/topics/standards/html
- What Works Clearinghouse (for evidence-based methods to improve teaching and learning)—www.w-w-c.org

Jeffery P. Braden, PhD, is a Professor in the School Psychology Program at North Carolina State University, and previously directed school psychology programs at the University of Wisconsin, San Jose State University, and the University of Florida. Jennifer L. Schroeder, PhD, NCSP, is a school psychologist in the Dallas Independent School District, Dallas, TX.

© 2004 National Association of School Psychologists, 4340 East West Highway, Suite 402, Bethesda, MD 20814—(301) 657-0270.



The National Association of School Psychologists (NASP) offers a wide variety of free or low cost online resources to parents, teachers, and others working with children and youth through the NASP website www.nasponline.org

and the NASP Center for Children & Families website www.naspcenter.org. Or use the direct links below to access information that can help you improve outcomes for the children and youth in your care.

About School Psychology—Downloadable brochures, FAQs, and facts about training, practice, and career choices for the profession.
www.nasponline.org/about_nasp/spsych.html

Crisis Resources—Handouts, fact sheets, and links regarding crisis prevention/intervention, coping with trauma, suicide prevention, and school safety.
www.nasponline.org/crisisresources

Culturally Competent Practice—Materials and resources promoting culturally competent assessment and intervention, minority recruitment, and issues related to cultural diversity and tolerance.
www.nasponline.org/culturalcompetence

En Español—Parent handouts and materials translated into Spanish. www.naspcenter.org/espanol/

IDEA Information—Information, resources, and advocacy tools regarding IDEA policy and practical implementation.
www.nasponline.org/advocacy/IDEAinformation.html

Information for Educators—Handouts, articles, and other resources on a variety of topics.
www.naspcenter.org/teachers/teachers.html

Information for Parents—Handouts and other resources a variety of topics.
www.naspcenter.org/parents/parents.html

Links to State Associations—Easy access to state association websites.
www.nasponline.org/information/links_state_orgs.html

NASP Books & Publications Store—Review tables of contents and chapters of NASP bestsellers.
www.nasponline.org/bestsellers
Order online. www.nasponline.org/store

Position Papers—Official NASP policy positions on key issues.
www.nasponline.org/information/position_paper.html

Success in School/Skills for Life—Parent handouts that can be posted on your school's website.
www.naspcenter.org/resourcekit

