



Accountability Issues and Reauthorization of the Elementary and Secondary Education Act

-name redacted-

Specialist in Education Policy

February 7, 2011

Congressional Research Service

7-....

www.crs.gov

R41533

CRS Report for Congress

Prepared for Members and Committees of Congress

Summary

Federal policies aiming to improve the effectiveness of schools have historically focused largely on inputs, such as supporting teacher professional development, class-size reduction, and compensatory programs or services for disadvantaged students. Over the last two decades, however, interest in developing federal policies that focus on student outcomes has increased. Most recently, the enactment of the No Child Left Behind Act of 2001 (NCLB; P.L. 107-110), which amended and reauthorized the Elementary and Secondary Education Act (ESEA), marked a dramatic expansion of the federal government's role in supporting standards-based instruction and test-based accountability, thereby increasing the federal government's involvement in decisions that directly affect teaching and learning.

As states and local educational agencies (LEAs) have implemented the federal accountability requirements, numerous issues have arisen that may be addressed during ESEA reauthorization. Among these issues are those pertaining to the comparability of data across states and the development of state accountability systems of varying degrees of rigor, the one-size-fits-all set of consequences applied to schools that fail to make adequate yearly progress (AYP), the narrowing of the focus of instruction at the school level, and the evaluation of teachers.

- **Commonality versus flexibility:** States have had the flexibility to select their own content and performance standards, as well as the assessments aligned with these standards. This has resulted in a different accountability system in each state, making it difficult to sum up where students are in terms of skills and knowledge and to gauge the net effect of the NCLB.
- **Absolute versus differentiated consequences:** Schools simply do or do not meet AYP standards, and there is generally no distinction between those that fail to meet only one or two required performance or participation thresholds to a marginal degree versus those that fail to meet numerous thresholds to a substantial extent.
- **Incentives to focus on proficiency:** Schools and LEAs are held accountable for the achievement of all student subgroups. They are not, however, held accountable for students at all levels of achievement. Because the goal of the current system is for 100% of students to become "proficient" by school year 2013-2014, schools and teachers may target instructional time and resources towards students who are nearing proficiency rather than distributing resources equally across students at all achievement levels.
- **Assessment and the narrowing of the curricular focus:** Because assessments are aligned with state content standards, there may be a risk that "teaching to the standards" becomes "teaching to the test." The practice of "teaching to the test"—whether intentional or unintentional—may narrow the curriculum.
- **Teacher evaluation and accountability:** The NCLB added a requirement that all teachers be highly qualified. Over time, however, the requirement has come to be seen by many as a minimum standard for entry into the profession and a growing body of research has revealed its underlying emphasis on teachers' credentials to be weakly correlated with student achievement. As such, the Administration has moved toward measuring teacher effectiveness based on student achievement, and promoted a focus on output-based accountability for teachers.

Contents

Introduction.....	1
Evolving Role of the Federal Government in Education-Related Accountability	1
Standards-Based Reform Movement.....	2
Goals 2000.....	3
The Improving America's Schools Act.....	5
The No Child Left Behind Act.....	7
Adequate Yearly Progress.....	8
Annual Measurable Objectives.....	9
Performance-Based Accountability: Consequences and Rewards	10
School Improvement, Corrective Action, and Restructuring	11
LEA Improvement and Corrective Actions	13
Rewards Based on Performance.....	13
Recent AYP Results for Schools and LEAs.....	14
Recent Performance-Based Accountability Results for Schools	15
Ability of ED to Affect Current Accountability Requirements	16
Issues Related to Implementation and Reauthorization.....	17
Commonality Versus Flexibility	18
Common Standards and Assessments.....	19
Absolute Versus Differentiated Consequences.....	21
Proficiency for All Students	23
Incentives to Focus on Proficiency.....	23
Assessment and the Narrowing of the Curricular Focus	24
Technical Quality of Assessments	26
Teacher Evaluation and Accountability.....	28

Tables

Table 1. ESEA Title I-A Performance-Based Accountability Stages for Schools and LEAs.....	10
Table 2. Number of States with Various Percentages of Public Schools and LEAs that Failed to Make Adequate Yearly Progress Based on AYP Determinations from the 2008-2009 School Year.....	14
Table 3. Number of States with Various Percentages of Title I-A Schools in Improvement, Corrective Action, or Restructuring During the 2009-2010 School Year Based on Adequate Yearly Progress Determinations from the 2008-2009 School Year	15
Table A-1. Percent of All Public Schools and LEAs that Did Not Make AYP: 2008-2009 School Year.....	31
Table A-2. Number of Title I-A Schools in Improvement, Corrective Action, or Restructuring During the 2009-2010 School Year Based on Adequate Yearly Progress Determinations from the 2008-2009 School Year	33

Appendixes

Appendix. Related Tables	31
--------------------------------	----

Contacts

Author Contact Information.....	36
Acknowledgments	36

Introduction

The effectiveness of the nation's schools is a concern at all levels of government. It is generally held that all students in elementary and secondary education should have access to quality public schools, providing all students with an opportunity to meet rigorous academic standards. Examining whether students are being held to well-defined academic standards and achievement at desired levels has become one of the primary foci of federal education policy in elementary and secondary schools.

Federal policies aiming to improve the effectiveness of schools have historically focused largely on inputs, such as supporting teacher professional development, class-size reduction, and compensatory programs or services for disadvantaged students. Over the last two decades, however, interest in developing federal policies that focus on student outcomes has increased. Most recently, the enactment of the No Child Left Behind Act of 2001 (NCLB; P.L. 107-110), which amended and reauthorized the Elementary and Secondary Education Act (ESEA), marked a dramatic expansion of the federal government's role in supporting standards-based instruction and test-based accountability, thereby increasing the federal government's involvement in decisions that directly affect teaching and learning.

The 112th Congress may act to reauthorize the ESEA. In approaching this task, Congress may grapple with several substantial issues related to the current accountability requirements included in the ESEA and how, or if, those requirements should be altered. For example, the following issues may receive attention: exploring ways to enhance consistency across state standards and assessments while ensuring sufficient flexibility for states with respect to these decisions, modifying consequences for failure to meet academic targets, incorporating outcomes-based accountability in teacher evaluations and compensation systems, and examining the role of the federal government in improving public elementary and secondary education.

The first part of this report examines the federal government's evolving role in supporting state and local accountability for elementary and secondary education and discusses how this role may continue to evolve or change during reauthorization. The discussion begins with the emergence of this role from the standards-based reform movement that gained traction in the 1980s and the resulting changes made to the ESEA during the 1994 reauthorization. This is followed by a detailed examination of the accountability requirements included in the NCLB. The discussion also examines the authority of the Secretary of Education (hereafter referred to as the Secretary) to make changes to the ESEA that some may view as effectively reauthorizing at least a portion of the ESEA. The second part of the report focuses on broad NCLB implementation issues that have arisen and how the federal role in elementary and secondary education may evolve or change in response during ESEA reauthorization.

Evolving Role of the Federal Government in Education-Related Accountability

Historically, education has been the domain of state and local governments. Since the initial enactment of the Elementary and Secondary Education Act in 1965 (P.L. 89-10), federal education policy has focused on equity. For example, since its inception the ESEA has focused on serving disadvantaged students and on meeting their academic needs to close the achievement

gap that exists between disadvantaged students and their more advantaged peers. Other federal education policies that were enacted after ESEA, such as the Individuals with Disabilities Education Act and Title IX, have also focused on providing equity. That is, federal education efforts in recent decades have been geared at pushing states and LEAs to provide equal educational opportunities for all students. The role of the federal government with respect to elementary and secondary education, however, began to change substantially in the early 1990s, as a focus on standards and assessments for all students was added to the traditional focus on equity. This section examines the evolution of this changing role beginning with the standards-based reform movement through the Improving America's Schools Act (IASA; P.L. 103-382), which reauthorized the ESEA in 1994. Subsequent sections of the report provide an in-depth analysis of the requirements of the NCLB and associated issues.

Standards-Based Reform Movement

During the 1980s, waves of education reform swept across the nation. The first wave included activities that were mostly characterized as “top-down” reforms, meaning that they were imposed on schools primarily from the state level.¹ These included, among other activities, increased requirements for high school graduation, attempts to lengthen the school day or year, increased testing of students' basic skills, and establishment of higher standards for entry or continuation in teaching. Generally, this initial wave of reform is considered to have had little positive effect on student achievement.²

Prompted in part by the limited success of the initial wave of the 1980s reforms, policymakers pursued other approaches to reform. The second wave of reform focused on empowerment, variation, and “bottom-up” decision making where policymakers sought to achieve the conditions necessary for educational success in individual schools. Some of these efforts focused on schools as the unit of change, calling for such action as greater school-site autonomy, increased competition among schools, and a greater role for parents in making educational decisions for their children.

In several states (e.g., Kentucky), a third approach that melded elements of top-down and bottom-up strategies was initiated. This course of action was premised on the need for system-wide reform, addressing the needs and problems at all levels of the system in a coherent and integrated fashion. At the heart of this reform was the premise that relevant reform action should be guided by educational goals and standards at each level of educational governance—particularly, state and local—and that reform action should occur simultaneously at these levels.

According to the Consortium for Policy Research in Education (CPRE), “[m]ore state activity aimed at improving public education took place in the 1980s than ever before.”³ This observation was based on the number of education-related bills that were introduced at the state level, increased state aid for education, and the number of state-level education-related task forces and commissions. This increase in state activity may be attributable, in part, to improved state

¹ Consortium for Policy Research in Education, *Putting the Pieces Together: Systemic School Reform*, RB-06-4/91, 1991, http://www.cpre.org/images/stories/cpre_pdfs/rb06.pdf.

² See, for example, *Education Reform in the '90s*, ed. Chester E. Finn, Jr. and Theodor Rebarber (1992), p. xi.

³ Center for Policy Research in Education, *State Education Reform in the 1980s*, RB-03-11/89, 1989, http://www.cpre.org/images/stories/cpre_pdfs_rb03.pdf.

policymaking capacities and growing state tax bases, as well as the publication of *A Nation At Risk*, which decried the condition of education in the nation.⁴

Not only were many states actively engaged in the standards-based reform movement, an event held in 1989 that brought together President George H. W. Bush and state governors, including then-Governor Bill Clinton of Arkansas, would mark a turning point in the federal government's involvement in the standards-based reform movement. At the first national education summit conference in the country's history, President Bush and the nation's governors endorsed the establishment of National Education Goals that would provide a framework for restructuring the federal role in elementary and secondary education. In the following year, they announced six goals. These policymakers agreed that setting goals would be an empty gesture unless progress toward the goals was measured and reported to the nation. To that end, they concluded that there should be standards defining what students need to know in each subject area, standards for performance in those subjects, and assessments to measure student performance. They also created the National Education Goals Panel (NEGP) to monitor national and state progress toward the goals.

Goals 2000

The groundwork laid at the first national education summit coalesced in the enactment of the Goals 2000: Educate America Act (P.L. 103-227)⁵ in 1994. Goals 2000 represented a new, structured federal role in the education reform movement that was occurring at the state and local levels. The first three titles of P.L. 103-227 codified eight National Education Goals to be achieved by 2000 (Title I), created a National Education Goals Panel to monitor progress toward the goals (Title II), and authorized a standards-based reform state grant program (Title III). Title I of Goals 2000 enacted into law the six goals originally adopted by President Bush and the governors in 1990 and added two additional goals (the fourth and eighth goals listed below). These goals provided that by the year 2000,

1. all children will start school ready to learn;
2. the high school graduation rate will be at least 90%;
3. students will master a challenging curriculum at grades 4, 8, and 12;
4. teachers will have access to professional development opportunities;
5. U.S. students will be the first in the world in science and math achievement;
6. all adults will be literate;
7. schools will be free of drugs, violence, and firearms; and
8. every school will promote parental involvement in education.

Three kinds of education standards were addressed initially by Goals 2000—curriculum content standards, student performance standards, and Opportunity-to-Learn standards.⁶

⁴ National Commission on Excellence in Education, *A Nation at Risk: The Imperative for Educational Reform*, April 1983, <http://www2.ed.gov/pubs/NatAtRisk/index.html>.

⁵ Goals 2000 includes all titles of P.L. 103-227 except Titles V and IX.

⁶ Opportunity-to-Learn (OTL) standards describe the resources needed to give all students an opportunity to acquire the knowledge and skills included in the content standards. The Omnibus Consolidated Rescissions and Appropriations Act (continued...)

To provide national oversight and direction of the standards effort, Title II authorized the establishment of two entities—NEGP (which was created prior to the enactment of Goals 2000) and the National Education Standards and Improvement Council (NESIC). While NESIC was never established,⁷ NEGP was charged with reporting annually on state and local progress toward the National Education Goals, reporting on promising or effective measures taken to achieve the goals, and assisting in the creation of a national consensus around the reforms needed to achieve the goals.

Title III authorized a formula grant program to support state standards-based reform by financing the development of state reform plans addressing such elements as teaching, learning, assessments, parental and community support for reform, and education management. Goals 2000 required that a state plan include processes for developing content and performance standards, developing state assessments, aligning curricula and assessments with state content and performance standards, and providing teachers with information about the state standards.

Also authorized by Title III were two waiver authorities. The first permitted the Secretary to waive various regulations of several major U.S. Department of Education (ED) programs for periods not to exceed four years, if requested by individual states or LEAs. To receive a waiver, states or LEAs had to be receiving systemic reform funds or be undertaking comparable reform activities. Under the second waiver authority, the Education Flexibility Partnership Demonstration Act (Ed-Flex), not more than six states⁸ could receive authority to determine, on their own, which statutory and regulatory requirements to waive in relation to several major ED programs.⁹

The Goals 2000 legislation triggered widespread controversy at the state and local level. The primary concerns arose over the proper federal role in education. Traditional federal education programs focused on special populations of students, particularly those that states and localities would not, or could not, adequately serve. In contrast, Goals 2000 targeted federal funds and support on some of the central functions of state and local educational systems (e.g., standards for curriculum and student performance) and to the needs of the entire student population. Debate on these issues centered largely on whether the legislation was part of a federal effort to centralize and standardize education.

(...continued)

of 1996 (P.L. 104-134) eliminated the requirement that states develop OTL standards or strategies as a condition of receiving Goals 2000 funds. It also repealed grant programs to develop national OTL standards and state assessments.

⁷ P.L. 104-134 repealed the authority to establish the NESIC. If it had been funded, NESIC would have certified voluntary national standards for curriculum content, pupil performance, and OTL, as well as voluntarily submitted state standards.

⁸ Participation was originally limited to six states, but P.L. 104-134 increased the participation limit to 12 states.

⁹ P.L. 106-25, the Education Flexibility Partnership Act, authorized Ed-Flex independent of the Goals 2000 legislation and allowed any state to participate. Currently 10 states have Ed-Flex authority: Colorado, Delaware, Kansas, Massachusetts, Maryland, North Carolina, Oregon, Pennsylvania, Texas, and Vermont. For more information, see CRS Report RL31583, *K-12 Education: Special Forms of Flexibility in the Administration of Federal Aid Programs*, by (name redacted).

The Improving America's Schools Act

Following directly on the heels of Goals 2000 was reauthorization of the ESEA. In reauthorizing the ESEA in 1994 through the Improving America's Schools Act (IASA; P.L. 103-382), Congress continued the federal focus on standards-based reform by inserting provisions related to standards and assessments in Title I-A. Title I-A authorizes federal grants to LEAs to provide supplementary educational and related services to low-achieving children attending schools with concentrations of students from low-income families. Title I-A is currently the largest federal elementary and secondary education program; it is funded at \$14 billion and accounts for over half of the \$25 billion appropriated for ESEA overall.

The IASA attempted to raise the instructional standards of Title I programs, and the academic expectations for participating students, by tying Title I instruction to state-selected content and performance standards. Further, IASA attempted to make Title I tests and evaluations more meaningful and less time consuming by using state-developed or state-adopted assessments, tied to the content standard, for determining the effectiveness of Title I programs. This measure of effectiveness was based on whether schools and LEAs were making "adequate yearly progress" (AYP) toward meeting the content and performance standards. The results on these assessments were also to become the basis for implementing program improvement requirements, including financial rewards to "distinguished" schools and LEAs and corrective actions for "unsuccessful" ones.

As a condition of the receipt of Title I-A funds, states were required to develop or adopt content and performance standards, as well as assessments aligned with the standards. If a state had standards and assessments that were applicable to all LEAs and students statewide, then these standards and assessments had to be used for Title I programs and students.¹⁰ However, the Title I statutory requirements did not require states to have standards and assessments that were applicable to all students statewide.¹¹ Thus, in discussions of "state systems of standards and assessments" in reference to IASA, it should be kept in mind that the standards and assessments were not necessarily uniform for all students statewide.

Under IASA, states were required to adopt standards and assessments in the subjects of reading/language arts (hereafter referred to as reading) and mathematics at three grade levels—at least once in each of the grade ranges of 3-5, 6-9, and 10-12. More specifically, in order to receive Title I-A funds, states were required to develop or adopt curriculum content standards, as well as academic achievement standards and assessments tied to the standards. The performance standards were required to establish three performance levels for all students—advanced, proficient, and partially proficient. States were also required to develop assessment systems that included all students in the grades being assessed who had attended schools in the LEA for at least one year. Limited English proficient (LEP) students were to be assessed, to the extent practicable, in the language and form most likely to yield accurate and reliable information on what they know and can do. "Reasonable" adaptations and accommodations had to be provided

¹⁰ If a state had developed content and performance standards under Title III of Goals 2000, or under a similar procedure, then those standards had to be used for Title I programs. The authorization for Title III of Goals 2000 expired at the end of FY2000.

¹¹ If a state did not have such statewide standards or assessments (as was the case for Iowa and Nebraska), then one of two alternative strategies could be adopted: (1) statewide standards applicable only to Title I students and programs, or (2) a mix of state and local, or local only, standards and assessments that were subject to state guidelines to ensure that they would meet the requirements of Title I.

for students with “diverse learning needs,” including LEP students and students with disabilities, where such adaptations and accommodations were necessary to measure the achievement of those students relative to state standards. The assessment system had to be capable of producing results for each state, LEA, and school that could be disaggregated by gender, major racial and ethnic groups, English proficiency status, migrant status, disability status, and economically disadvantaged status. These data were to be included in annual school profiles.

Several statutory constraints were placed on the authority of the Secretary in enforcing the standards and assessment requirements. For example, state standards did not have to be linked to national standards that might be developed by any organization or entity. No state could be required to submit its content or performance standards or assessments to the U.S. Department of Education (ED). ED was prohibited from disapproving a state plan on the basis of specific content or performance standards or assessments. In addition, nothing in the act authorized ED to “mandate, direct, or control” a state’s, LEA’s, or school’s standards, curricula, assessments, or program of instruction as a condition for receipt of Title I funds.

States were given several years to meet the IASA requirements, in part because states were at widely varying stages of the process of developing instructional goals, curriculum frameworks, and assessment systems that were aligned with these goals and frameworks. States were given until the beginning of the 1997-1998 school year to develop or adopt content and performance standards for reading and mathematics. Assessments that were aligned with these standards had to be developed or adopted by the start of the 2000-2001 school year.¹² The full system of standards and assessments was not required to be in place until the 2000-2001 school year. Under IASA, state standards developed under Title III of Goals 2000 were considered to meet the Title I-A standards requirements under IASA.

Despite the increased federal emphasis on the development and implementation of state standards and assessments for all students, about half of the states failed to meet the 1997-1998 school year deadline included in Title I-A for the development of content standards. By March 2000, however, all but two states had adopted the required content standards, but only about half of the states had approved performance standards.

It is important to note that IASA did not mark the beginning of federal efforts to hold states and LEAs accountable for student performance under Title I-A. Since enactment of the ESEA in 1965, the ESEA has included requirements that states and LEAs report on the academic progress of students being served under Title I-A. Subsequent evaluations of Title I found that LEAs were not effectively implementing testing systems, were using tests to draw inferences that were inappropriate based on the tests, and were measuring only basic skills.¹³ That is, LEAs essentially had low expectations for disadvantaged students, rather than high expectations for all students. The standards and assessment provisions enacted through IASA addressed three long-standing and widely shared concerns about Title I programs. In addition to believing that Title I programs had not been sufficiently challenging academically, perpetuating low expectations for the achievement of participating students, many analysts also felt that Title I programs had not been well integrated with the “regular” instructional programs of participants and had required

¹² The Secretary was authorized to provide a one-year extension of the deadline for assessments for the correction of problems identified in the field testing of assessments.

¹³ Paul Manna, *Collision Course: Federal Education Policy Meets State and Local Realities* (Washington, DC: CQ Press, 2010).

extensive student testing that was of little instructional or diagnostic value and was not linked to the curriculum to which the students were exposed.

The No Child Left Behind Act

Many of the underpinnings of the accountability requirements included in Title I-A of the No Child Left Behind Act of 2001 (NCLB; P.L. 107-110) are evident in the accountability requirements included in the IASA; in some cases, the requirements were substantially expanded. For example, in addition to the IASA's requirement for states to implement standards and assessments in reading and mathematics at three grade levels, the NCLB required states participating in ESEA Title I-A to

- develop and adopt standards and assessments in the subjects of mathematics and reading in *each* of grades 3-8 by the end of the 2005-2006 school year, assuming certain minimum levels of annual federal funding were provided for state assessment grants;
- adopt standards in science (at three grade levels—grades 3-5, 6-9, and 10-12) by the end of the 2005-2006 school year; and
- adopt assessments in science (at three grade levels) by the end of the 2007-2008 school year.¹⁴

The academic achievement standards must include at least three levels of performance: partially proficient (basic), proficient, and advanced. The same academic content and achievement standards must apply to all students.¹⁵ The assessments must be aligned with the state's academic content and achievement standards.

State accountability systems must be based on the academic standards and assessments required above, be the same accountability system for all public schools (except that public schools and LEAs that do not receive Title I-A funds are not subject to outcome accountability requirements), and incorporate rewards and sanctions based on student performance. To the extent practicable, LEP students are to be assessed in the language and form most likely to yield accurate and reliable information on what they know and can do in academic content areas (in subjects other than English itself). However, students who have attended schools in the United States (excluding Puerto Rico) for three or more consecutive school years are to be assessed in English.¹⁶ In addition, "reasonable" adaptations and accommodations are to be provided for students with disabilities, consistent with the provisions of the IDEA.¹⁷

¹⁴ As with IASA, if no agency or entity in a state has authority to establish statewide standards or assessments (as is generally assumed to be the case for Iowa and Nebraska), then the state may adopt either (1) statewide standards and assessments applicable only to Title I-A students and programs, or (2) a policy providing that each LEA receiving Title I-A grants will adopt standards and assessments that meet the requirements of Title I-A and are applicable to all students served by each such LEA.

¹⁵ One exception to this provision involves students with disabilities. For more information, see CRS Report R40701, *Alternate Assessments for Students with Disabilities*, by (name redacted).

¹⁶ LEAs may continue to administer assessments to pupils in non-English languages for up to a total of five years if, on a case-by-case basis, they determine that this would likely yield more accurate information on what the students know and can do.

¹⁷ For further information on this and related topics, see CRS Report RL32913, *The Individuals with Disabilities* (continued...)

Similar to IASA, several statutory constraints have been placed on the authority of the Secretary to enforce these standard and assessment requirements. First, the ESEA states that nothing in Title I shall be construed to authorize any federal official or agency to “mandate, direct, or control a State, local educational agency, or school’s specific instructional content, academic achievement standards and assessments, curriculum, or program of instruction.” Second, states may not be required to submit their standards to the Secretary or to have their content or achievement standards approved or certified by the federal government in order to receive funds under the ESEA, other than the (limited) review necessary in order to determine whether the state meets the Title I-A technical requirements. Finally, no state plan may be disapproved by ED on the basis of specific content or achievement standards, or assessment items or instruments.

Adequate Yearly Progress

A key concept embodied in the outcome accountability requirements of the NCLB—that of AYP for schools, LEAs, and (with much less emphasis) states overall¹⁸—expands on earlier provisions included in the IASA. The primary purpose of AYP requirements is to serve as the basis for identifying schools and LEAs where performance is inadequate, so that these inadequacies may be addressed, first through provision of increased support and opportunities for families to exercise choice to transfer to another school or obtain supplemental educational services (SES), and ultimately through a series of more substantial consequences (described in a later section of this report, “Performance-Based Accountability”). These actions are to be taken with respect to schools or LEAs that fail to meet AYP for two consecutive years or more.

AYP standards under the NCLB must be applied to *all* public schools, LEAs, and to states overall, if a state chooses to receive Title I-A grants.¹⁹ However, consequences for failing to meet AYP standards for two consecutive years or more need only be applied to schools and LEAs participating in Title I-A,²⁰ and there are no sanctions for states overall beyond potential identification and the provision of technical assistance.

AYP is determined based on three components: student academic achievement on the required state reading and mathematics assessments, with a focus on the percentage of students scoring at the proficient level or higher;²¹ 95% student participation rates in assessments by all students and for any subgroup for which data are disaggregated; and performance on another academic indicator, which must be graduation rates for high schools.²² Schools or LEAs meet AYP standards only if they meet the required threshold levels of performance on all three indicators for the all-students group and any subgroup for which data are disaggregated. AYP calculations based

(...continued)

Education Act (IDEA): Interactions with Selected Provisions of the No Child Left Behind Act (NCLB), by (name redacted) and (name redacted).

¹⁸ For more information on this topic, see CRS Report RL32495, *Adequate Yearly Progress (AYP): Implementation of the No Child Left Behind Act*, by (name redacted); and CRS Report RL33731, *Education for the Disadvantaged: Reauthorization Issues for ESEA Title I-A Under the No Child Left Behind Act*, by (name redacted).

¹⁹ Currently, all states receive Title I-A funds.

²⁰ States are encouraged to apply these consequences to *all* public schools and LEAs, but are not required to do so. State practices vary on this point.

²¹ Results from the science assessments do not have to be included in the AYP calculation.

²² States select the other academic indicator for elementary and middle schools. Generally, states choose to use attendance rates.

on assessment scores must be disaggregated; that is, they must be determined separately and specifically not only for all students but also for several demographic groups of students within each school, LEA, and state. The specified demographic groups (often referred to as subgroups) are economically disadvantaged students, LEP students, students with disabilities, and students in major racial and ethnic groups.²³

It is generally acknowledged that one of the most positive contributions of the NCLB requirements has been the reporting of data by student subgroups, thereby making it possible to see how disadvantaged students are performing relative to other students. The results highlight substantial achievement gaps between minority students and white students, LEP students and non-LEP students, students with and without disabilities, and economically disadvantaged students and their more economically advantaged peers. Having these results echoed nationwide has spurred and may continue to spur attention at the state and local levels on meeting the academic needs of underserved students, which has been a primary goal of ESEA since its inception. Changes to the ESEA over the past two decades have also helped to ensure that all students are held to the same academically challenging standards rather than having lower academic expectations for students served by the Title I-A program. What is unclear, however, is whether the law has actually improved student achievement for the majority of disadvantaged students and whether the federal government has employed policy levers that can bring about this change through the ESEA.

Annual Measurable Objectives

States must develop annual measurable objectives (AMOs) that are established separately for reading and mathematics assessments, are the same for all schools and LEAs, identify a single minimum percentage of students who must meet or exceed the proficient level on the assessments that applies to the all students group and each subgroup for which data are disaggregated, and must ensure that all students will meet or exceed the state's proficient level of achievement on the assessments based on a timeline established by the state. The timeline must incorporate concrete movement toward meeting an *ultimate goal* of all students reaching a proficient or higher level of achievement by the end of the 2013-2014 school year. This was adopted in response to pre-NCLB AYP requirements in most states that required little or no net improvement in student performance over time.

While NCLB substantially increased federal involvement in student assessment and learning, an argument could be made that this involvement stems from the traditional federal education policy

²³ Several constraints apply to the consideration of these student groups in AYP calculations. First, student groups need not be considered in cases where their number is so relatively small that achievement results would not be statistically significant or the identity of individual pupils might be divulged. The selection of the minimum number ("n") of students in a group for the group to be considered in AYP determinations has been left largely to state discretion. Since the same minimum group size policies are applied to schools and to LEAs overall, groups that are too small to be separately considered for individual schools often meet the minimum group size threshold at the LEA level. Second, it has been left to the states to define the "major racial and ethnic groups" on the basis of which AYP must be calculated. And third, pupils who have not attended the same school for a full year need not be considered in determining AYP at the school level, although they are still to be included in LEA and state AYP determinations, if they attended schools in the same LEA or state for the full academic year. A number of special rules, which have evolved over time, apply to two of the disaggregated pupil groups: LEP pupils and pupils with disabilities. For a detailed discussion of these, see CRS Report RL32495, *Adequate Yearly Progress (AYP): Implementation of the No Child Left Behind Act*, by (name redacted); and CRS Report RL32913, *The Individuals with Disabilities Education Act (IDEA): Interactions with Selected Provisions of the No Child Left Behind Act (NCLB)*, by (name redacted) and (name redacted).

on equity. By requiring each state to ensure that all students make progress toward the ultimate goal of proficiency, that all students are held to the same content and performance standards, and that results are disaggregated and publicly reported for historically disadvantaged student groups, the federal requirements have pressured states into developing one set of expectations for all students and forced them to acknowledge when specific groups of students are not meeting those expectations. Without this federal involvement, it is impossible to know whether states would have taken similar actions. However, as discussed in a subsequent section, the expectations for students across states are anything but consistent.

Performance-Based Accountability: Consequences and Rewards

The NCLB requires states to identify LEAs, and LEAs to identify schools, that fail to meet state AYP standards for two consecutive years for program improvement, and to take a variety of further actions with respect to schools or LEAs that fail to meet AYP standards for additional years after being identified for improvement.²⁴ Although states are encouraged to establish unitary accountability systems affecting all public schools, and some states do so to varying degrees, the Title I-A statute only requires them to apply performance-based accountability measures to schools and LEAs that receive Title I-A funds, not all schools and LEAs.²⁵ The “accountability stages” are depicted in **Table 1**.

Table 1. ESEA Title I-A Performance-Based Accountability Stages for Schools and LEAs

Cumulative Years Not Making AYP	Accountability Stage	
	School	LEA
1	N/A	N/A
2	School improvement	LEA improvement ^a
3	Second year of school improvement ^b	LEA improvement ^a
4	Corrective action ^b	Corrective action ^b
5	Plan for restructuring ^b	Corrective action ^b
6	Implement restructuring ^b	Corrective action ^b

Source: Table prepared by CRS based on analysis of the requirements of ESEA, Section 1116.

Note: N/A—Not Applicable.

- a. State educational agencies (SEAs) may implement corrective action for an LEA identified for LEA improvement.
- b. Accountability requirements associated with the second year of school improvement, corrective action, and restructuring may be delayed for up to one year for a school or LEA if it makes AYP for one year, or if its failure to make AYP is due to a natural disaster or a significant decline in financial resources.

²⁴ ED is required to establish a peer review process to evaluate whether states overall have met their statewide AYP goals, beginning after the third year of implementation of the NCLB. States that fail to meet their goals are to be listed in an annual report to Congress, and technical assistance is to be provided to states that fail to meet their goals for two consecutive years or more. For further information on this topic, see CRS Report RL33371, *K-12 Education: Implementation Status of the No Child Left Behind Act of 2001 (P.L. 107-110)*, coordinated by (name redacted).

²⁵ Over 90% of all LEAs, and approximately three-fifths of all public schools, participate in the Title I-A program.

School Improvement, Corrective Action, and Restructuring

Title I-A schools that fail to meet AYP standards for two consecutive years must be identified for program improvement. When Title I-A schools do not make AYP for two or more consecutive years, they become subject to a range of increasingly severe performance-based accountability requirements, which are coupled with technical assistance provided by the LEA.

School Improvement

After not making AYP for two consecutive years, a Title I-A school is identified for school improvement. Being designated for school improvement carries with it the requirement to develop or revise a school plan designed to result in the improvement of the school. LEAs are required to provide schools within their jurisdictions with technical assistance in the design and implementation of school improvement plans. Schools identified for improvement must use at least 10% of their Title I-A funding for professional development.

All students attending Title I-A schools identified for school improvement must be offered public school choice—the opportunity to transfer to another public school within the same LEA.²⁶ Under public school choice, students must be afforded the opportunity to choose from among two or more schools, located within the same LEA, that have not been identified for school improvement, corrective action, or restructuring, and that also have not been identified as persistently dangerous schools.²⁷ LEAs are required to provide students who transfer to different schools with transportation and must give priority in choosing schools to the lowest-achieving children from low-income families. LEAs may not use lack of capacity as a reason for denying students the opportunity to transfer to a school of choice.²⁸ In instances where there are no eligible schools in the student's LEA, LEAs are encouraged to enter into cooperative agreements with surrounding LEAs to enable students to transfer to an eligible public school.

If a school does not make AYP for another year after being identified for an initial year of school improvement, it must be identified for a second year of school improvement. All students attending a school identified for a second year of school improvement must continue to be offered the option of attending another eligible public school within the same LEA. In addition, students from low-income families who continue to attend the school must be offered the opportunity to receive supplemental educational services (SES).²⁹ SES are educational activities, such as tutoring, that are provided outside of normal school hours and which are designed to augment or enhance the educational services provided during regular periods of instruction. Supplemental educational services may be provided by a non-profit entity, a for-profit entity, or the LEA, unless such services are determined by the state educational agency (SEA) to be unavailable in the local area.³⁰ The SEA is required to maintain a list of approved SES providers (including those offering

²⁶ For further information on public school choice, see CRS Report RL33506, *School Choice Under the ESEA: Programs and Requirements*, by (name redacted).

²⁷ For more information on persistently dangerous schools, see CRS Report RL33371, *K-12 Education: Implementation Status of the No Child Left Behind Act of 2001 (P.L. 107-110)*, coordinated by (name redacted).

²⁸ 34 CFR 200.44(d).

²⁹ For further information on supplemental educational services, see CRS Report RL31329, *Supplemental Educational Services for Children from Low-Income Families Under ESEA Title I-A*, by (name redacted).

³⁰ Schools identified for improvement, corrective action, or restructuring, and LEAs identified for improvement or corrective action, lose their eligibility to supplemental educational services providers.

services through distance learning) from which parents can select. LEAs may be required to expend up to an amount equal to 20% of their Title I-A grants on transportation for public school choice and supplemental educational services combined.³¹

Corrective Action

If a school fails to make AYP for a total of two years after being identified for school improvement, it must be identified for corrective action by the end of the school year. For schools identified for corrective action, LEAs must continue to provide technical assistance, offer public school choice and supplemental educational services, and implement one of the following corrective actions: replacing school staff relevant to the school not making AYP, implementing a new curriculum, limiting management authority at the school level, appointing an expert advisor to assist in implementing the school improvement plan, extending the school year or the school day, or restructuring the school's internal organization.

Restructuring

If a school does not make AYP for a third year after being identified for school improvement, by the end of the school year the LEA must begin to plan for restructuring while continuing to implement the requirements of corrective action. Restructuring of the school must involve implementation of some form of alternative governance structure, such as reopening the school as a charter school, replacing all or most of the school staff, contracting with an education management organization to operate the school, or turning the school over to the SEA. If an additional year passes without the school making AYP, the LEA must implement restructuring of the school.

Any of the sanctions described above may be delayed for up to one year if the school makes AYP for a single year, or if the school's failure to make AYP is due to unforeseen circumstances, such as a natural disaster or a significant decline in financial resources of the LEA or school. Schools that make AYP for two consecutive years may no longer be identified for school improvement nor be subject to the sanctions associated with school improvement, corrective action, or restructuring.

In addition, in 2008 ED announced a pilot program under which the Secretary would grant waivers to up to 10 states proposing to implement alternative performance-based accountability policies incorporating differentiated consequences.³² Under these policies, states could distinguish among schools identified for improvement, focusing resources upon, and applying the most significant consequences to, schools with the lowest performance levels. Among the nine states initially approved for the pilot program,³³ the state accountability plans have three characteristics in common: (1) placement of schools identified for improvement into two or more categories, particularly a highest-priority group on which the most substantial consequences and improvement resources would be focused; (2) some adjustment of the order and/or severity of

³¹ More specifically, LEAs are to use an amount equal to 5% of their Title I-A grant for public school choice and transportation costs, 5% for SES, and up to an additional 10% for either, to the extent needed. These funds may be taken from the LEA's Title I-A grant or from other sources.

³² See <http://www.ed.gov/news/pressreleases/2008/03/03182008.html>.

³³ The nine states are Arkansas, Florida, Georgia, Illinois, Indiana, Louisiana, Maryland, New York, and Ohio.

consequences for schools placed into different improvement categories, particularly with respect to school choice and SES; and (3) in many cases, narrowing of certain consequences or actions to focus more specifically on pupil groups with the lowest levels of performance.³⁴

As the 2013-2014 deadline for all students to be proficient in reading/language arts and mathematics approaches, the number of schools that fail to make AYP will almost certainly increase as the percentage of students required to be at the proficient level increases. As the number of schools failing to make AYP grows, states and LEAs may become increasingly taxed with respect to their ability to help all of the schools.

LEA Improvement and Corrective Actions

Procedures analogous to those for schools apply to LEAs that receive Title I-A grants and fail to meet AYP requirements. While states are encouraged to implement unitary accountability systems applicable to all students and schools, states may choose to base decisions regarding LEA status and corrective actions only on the Title I-A schools in each LEA. Further, as noted earlier, identification as needing improvement and corrective actions need only be taken with respect to LEAs that receive Title I-A grants, although this includes over 90% of all LEAs nationwide.

LEAs that fail to meet state AYP standards for two consecutive years are to be identified as needing improvement. Technical assistance, “based on scientifically based research,” is to be provided to the LEA by the SEA, and parents of students served by the LEA are to be notified that it has been identified as needing improvement.

SEAs are to take corrective action with respect to LEAs that fail to meet state standards for a fourth year (two years of failing to meet AYP standards after having been identified for improvement without, in the meantime, meeting AYP standards for two consecutive years). Such corrective action is to include at least one of the following (at SEA discretion): reducing administrative funds or deferring program funds, implementing a new curriculum, replacing relevant LEA staff, removing specific schools from the jurisdiction of the LEA, appointing a receiver or trustee to administer the LEA, abolishing or restructuring the LEA, or authorizing students to transfer to higher-performing schools in another LEA (and providing transportation) in conjunction with at least one of these actions.

Rewards Based on Performance

Each state participating in ESEA Title I-A is required to establish an Academic Achievement Awards Program for purposes of making academic achievement awards to schools that have either significantly closed academic achievement gaps between student subgroups or exceeded their AYP requirements for two or more consecutive years. States may also give awards to LEAs that have exceeded their AYP requirements for two or more consecutive years. Under Academic Achievement Awards Programs, states may recognize and provide financial awards to teachers or principals in schools that have significantly closed the academic achievement gap or that have made AYP for two consecutive years. States may fund Academic Achievement Awards for

³⁴ The state plans may be found at <http://www.ed.gov/admins/lead/account/differentiatedaccountability/index.html>.

schools and LEAs by reserving up to 5% of any Title I-A funding that is in excess of the state's previous-year allocation.³⁵

Recent AYP Results for Schools and LEAs

Nationwide, 33.7% of all public schools and 36.3% of all LEAs failed to make AYP based on AYP determinations for the 2008-2009 school year.³⁶ **Table 1** provides data on the percentage of schools and LEAs failing to make adequately yearly progress, on the basis of 2008-2009 assessment results, as reported by ED, based on Consolidated State Performance Reports.³⁷ These data are based on all public schools and LEAs in each state, not just those participating in Title I-A.³⁸ Overall, 48 states, including the District of Columbia and Puerto Rico, had at least 10% of their public schools fail to make AYP based on AYP determinations for the 2008-2009 school year (**Table 2**). Ten states, however, had over 50% of their public schools fail to make AYP. At the LEA level, 38 states had at least 10% of their LEAs fail to make AYP, while 19 states had at least half of their LEAs fail to make AYP.³⁹ The percentage of public schools failing to make adequate yearly progress for 2008-2009 varied widely among the states, from 5% for Texas to 76.6% for Florida (**Table A-1**).⁴⁰ Among the states, there was even greater variation with respect to the percentage of LEAs making AYP. Two states—Maine and Wisconsin—reported that just under 1% of their LEAs failed to make adequate yearly progress, while 94.5% of the LEAs in West Virginia failed to meet AYP standards.

Table 2. Number of States with Various Percentages of Public Schools and LEAs that Failed to Make Adequate Yearly Progress Based on AYP Determinations from the 2008-2009 School Year

Entity	Number of States Where More Than 10% of Public Schools or LEAs Failed to Make AYP	Number of States Where More Than 25% of Public Schools or LEAs Failed to Make AYP	Number of States Where More Than 50% of Public Schools or LEAs Failed to Make AYP
Public schools	48	35	10
LEAs	38	28	19

³⁵ States may fund teacher and principal awards by reserving such sums as necessary from the amount received under ESEA Title II-A-1—Teacher and Principal Training and Recruiting Fund, Grants to States. Guidance on procedures for reserving funds for State Academic Achievement Awards Programs is available in U.S. Department of Education, Office of Elementary and Secondary Education, *Guidance: State Educational Agency Procedures for Adjusting Basic, Concentration, Targeted, and Education Finance Incentive Grant Allocations Determined by the U.S. Department of Education*, May 23, 2003, pp. 32-34.

³⁶ U.S. Department of Education, Education Data Express System. The original source of the data is the EDFacts/Consolidated State Performance Report, 2008-09, available online at <http://www2.ed.gov/admins/lead/account/consolidated/index.html>.

³⁷ See <http://www.ed.gov/admins/lead/account/consolidated/index.html>.

³⁸ Data are also available on Title I-A-recipient schools and LEAs that fail to make adequate yearly progress. However, in the aggregate, the results are quite similar. For 2008-2009, 35% of all public schools and 36% of all Title I-A schools were reported as failing to make AYP. For LEAs, 35% of all LEAs and 38% of all Title I-A LEAs failed to make AYP.

³⁹ It should be noted that neither Hawaii nor Puerto Rico reported LEA-level data. Both states have only one LEA. In both cases, the single LEA failed to make AYP based on determinations for the 2008-2009 school year. Therefore, 21 states had at least half of their LEAs fail to make AYP.

⁴⁰ In Puerto Rico, 93.5% of public schools failed to make AYP.

Source: Table prepared by CRS, November 24, 2010, based on data available from the U.S. Department of Education, Education Data Express System. The original source of the data is the EDFacts/Consolidated State Performance Report, 2008-09, available online at <http://www2.ed.gov/admins/lead/account/consolidated/index.html>.

Notes: Not all states reported data for all types of performance-based accountability sanctions. Therefore, due to missing values, national level data may not accurately represent national trends. The District of Columbia and Puerto Rico have been included as “states.”

Recent Performance-Based Accountability Results for Schools

As shown in **Table 3**, most states, including the District of Columbia and Puerto Rico, had at least 10% of their Title I-A schools in some form of performance-based accountability sanction during the 2009-2010 school year, and over half of the states have at least 20% of their Title I-A schools in improvement, corrective action, or restructuring.⁴¹ Overall, about 28% of Title I-A schools in the nation have been identified for improvement (**Table A-2**). However, the percentage of schools identified varies widely by state, ranging from about 3% of Title I-A schools in Oklahoma to about 85% of Title I-A schools in Georgia.

While the percentages of schools identified for some type of improvement vary across the states, so too does the percentage of students taking advantage of the public school choice and SES options.⁴² For example, during the 2008-2009 school year 2.7% of all students eligible for public school choice took advantage of this option.⁴³ However, the participation rate by state varied from less than 0.1% (e.g., District of Columbia) to nearly 80% in South Dakota. In addition, 15.6% of eligible students took advantage of the availability of SES.⁴⁴ As with the school choice option, the participation rate varied substantially by state, ranging from less than 1% in Montana to over 97% in the District of Columbia and 40.2% in Maryland.

Table 3. Number of States with Various Percentages of Title I-A Schools in Improvement, Corrective Action, or Restructuring During the 2009-2010 School Year Based on Adequate Yearly Progress Determinations from the 2008-2009 School Year

Performance-Based Accountability Sanction	Number of States with More than 10% of Their Title I-A Schools Subject to the Sanction	Number of States with More than 20% of Their Title I-A Schools Subject to the Sanction
School in improvement, year one	19	3
School in improvement, year two	9	0
Corrective action	4	0
Planning for restructuring	5	1

⁴¹ Data were not readily available for LEAs.

⁴² For more information on reasons why students may not be taking advantage of public school choice and SES options, see U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service, *State and Local Implementation of the No Child Left Behind Act: Volume VII—Title I School Choice and Supplemental Educational Services: Final Report*, Washington, DC, 2009, <http://www.ed.gov/rschstat/eval/choice/nclb-choice-ses-final/choice-ses-final.pdf>.

⁴³ Data obtained by CRS from the U.S. Department of Education, ED Data Express System, November 22, 2010.

⁴⁴ Data obtained by CRS from the U.S. Department of Education, ED Data System, November 22, 2010.

Performance-Based Accountability Sanction	Number of States with More than 10% of Their Title I-A Schools Subject to the Sanction	Number of States with More than 20% of Their Title I-A Schools Subject to the Sanction
Implementing restructuring	14	5
School subject to any of the performance-based accountability sanctions (i.e., improvement, corrective action, or restructuring)	42	29

Source: Table prepared by CRS, November 22, 2010, based on data available from the U.S. Department of Education, Education Data Express System. The original source of the data is the EDFacts/Consolidated State Performance Report, 2008-09, available online at <http://www2.ed.gov/admins/lead/account/consolidated/index.html>.

Notes: A Title I-A school is a school that received federal funds through the Title I-A program authorized by the Elementary and Secondary Education Act (ESEA). States are only required to apply the performance-based outcome accountability requirements to Title I-A schools that fail to make adequate yearly progress. The District of Columbia and Puerto Rico have been included as “states.” There are no data for Indiana because the state developed new improvement statuses under a differentiated accountability pilot that do not correspond with the existing improvement statuses. Not all states reported data for all types of performance-based accountability sanctions. Therefore, due to missing values national-level data may not accurately represent national trends.

Ability of ED to Affect Current Accountability Requirements

ED has the ability through regulatory and waiver authority to alter the current requirements of the ESEA. As part of its role, the Secretary has the authority to issue regulations that affect program implementation. Both Section 410 of the General Education Provisions Act (GEPA) and Section 9535 of the ESEA provide the Secretary with the authority to promulgate regulations related to programs administered by the Secretary. In addition to regulatory authority, Section 9401 grants the Secretary the authority to grant waivers of any statutory or regulatory requirement of the ESEA⁴⁵ for a SEA, LEA, Indian tribe, or school. Entities requesting waivers must meet specific requirements, such as describing for each school year the “specific, measurable education goals” in accordance with the Title I-A accountability requirements for the SEA (if applicable) and for each LEA, Indian tribe, or school that would be affected by the waiver and the methods that would be used to measure annual progress toward meeting such goals and outcomes.

Since the reauthorization of the ESEA in 2002, and particularly as Congress has worked toward reauthorizing the ESEA, the Secretary has used regulatory authority to make substantive changes to ESEA program requirements. For example, former Secretary of Education Spellings took the opportunity to expand federal regulation of ESEA programs as well as allow selected, substantial forms of flexibility through the use of waiver authority provided under Section 9401. New regulations published in October 2008 substantially expanded ESEA Title I-A requirements dealing with high school graduation rates, as used in AYP determinations, as well as school choice and supplemental educational services for students in schools identified for improvement.⁴⁶ At the same time, these regulations codify a previously initiated waiver program

⁴⁵ There are some limits placed on the Secretary’s ability to grant waivers of ESEA provisions. For example, requirements related to fiscal accountability, civil rights, parental participation, and the equitable participation of private school students may not be waived.

⁴⁶ The October 2008 regulations are available online at <http://www2.ed.gov/legislation/FedRegister/finrule/2008-4/102908a.html>.

allowing use of growth models of AYP, and a new waiver pilot allowing a limited number of states to implement differentiated consequences for schools that fail to make AYP.

With respect to the use of waiver authority, ED is required to publish an annual accounting of all waivers granted under Section 9401. The most recent announcement covers waivers granted during calendar year 2009.⁴⁷ During calendar year 2009, the Secretary issued 351 waivers. Over half of the waivers granted (196) were provided to LEAs and schools with respect to the treatment of their Title I-A funds granted under the American Recovery and Reinvestment Act (ARRA; P.L. 111-5).⁴⁸ The majority of the remaining waivers addressed ESEA-specific issues (as opposed to issues resulting from ARRA enactment) such as the “growth model” pilot and the “differentiated accountability” pilot.⁴⁹

Collectively, these actions reflect a trend toward influence over design elements of accountability requirements by Administration initiatives, as opposed to statutes. Using waiver authority, in particular, it is possible that the Secretary could act to expand upon the specific provisions for which he would be willing to grant waivers; thereby, further changing the accountability requirements included in Title I-A, regardless of whether Congress acts to reauthorize the ESEA. Given that a large percentage of schools and LEAs are already failing to make AYP, and a good deal more are expected to join those ranks as 100% proficiency is required by the end of the 2013-2014 school year, it has been suggested that the Secretary could act to make some adjustments to the NCLB accountability requirements through waiver authority if Congress does not act to reauthorize the ESEA.

Issues Related to Implementation and Reauthorization

The standards-based movement and accompanying test-based accountability movement were initiated at the state and local levels. It does not appear likely that states would abandon efforts to develop rigorous content standards and implement measures to determine whether students have learned the material included in these standards. Federal involvement in public K-12 education, however, is significantly more extensive than in the past, while the aggregate federal contribution to public K-12 education revenues remains relatively small (approximately 9%). In considering the ESEA for reauthorization, Congress may decide whether to continue this active federal strategy, perhaps expanding it further through increased assessment or other requirements, or alternatively decide to place tighter limits on the scope of federal involvement in state and local

⁴⁷ U.S. Department of Education, “Notice of Waivers Granted Under Section 9401 of the Elementary and Secondary Education Act of 1965, as Amended,” 75 *Federal Register* 56833-56856, September 16, 2010.

⁴⁸ For example, LEAs received waivers to exclude the amount Title I-A funds received under ARRA when calculating their obligation to spend their funds for public school choice and professional development and when calculating the per-pupil amount for SES, and to waive the carryover limitation more than once every three years. For more information, see U.S. Department of Education, “Notice of Waivers Granted Under Section 9401 of the Elementary and Secondary Education Act of 1965, as Amended,” 75 *Federal Register* 56833-56856, September 16, 2010.

⁴⁹ Under growth models, the achievement of the same pupils is tracked from year-to-year. This type of model is not explicitly mentioned in the NCLB statute; however, it is authorized in regulations promulgated by ED. Using waiver authority available to the Secretary under Section 9401, the Secretary is able to approve a state’s use of growth models. To date, 15 states (Alaska, Arizona, Arkansas, Colorado, Delaware, Florida, Iowa, Michigan, Minnesota, Missouri, North Carolina, Ohio, Pennsylvania, Tennessee, and Texas) have been approved to use growth models.

K-12 education systems. A hybrid approach might involve continued or expanded federal requirements regarding student outcomes combined with fewer requirements regarding the purposes for which federal grant funds can be used. Examples of the latter might include program consolidation, an expansion of current authority to transfer funds among ESEA programs, or policies offering increased flexibility in return for reaching specified levels of performance. It is also unclear what role common standards and assessments (discussed below) might play in an ESEA reauthorization effort. It is possible that a tradeoff could be made in this area as well. For example, states that adopt the common standards and assessments could receive increased flexibility in the consequences applied to schools and LEAs that fail to make AYP.

The remainder of the report highlights and examines several broad issues related to the implementation of the NCLB and possible issues to consider during the reauthorization of the ESEA. Several policy-related issues have arisen as NCLB requirements have been implemented. These include issues pertaining to the comparability of data across states and the development of state accountability systems of varying degrees of rigor, the one-size-fits all set of consequences applied to schools that fail to make AYP, the narrowing of the focus of instruction at the school level, and the evaluation of teachers.

Commonality Versus Flexibility

The accountability requirements included in the NCLB represented unprecedented federal involvement in school- and LEA-level accountability for student performance; however, the requirements are still broad enough that states were able to select their own content standards, and determine how those standards would be assessed (i.e., which tests would be used), and what would constitute a proficient level of performance on the tests. Thus, a byproduct (perhaps unintended) of NCLB enactment has been the creation of more than 50 different accountability systems across the nation. As state accountability systems differ so substantially, one of the tradeoffs in this approach is the lack of comparable data on student performance across states. Because states are using different standards and assessments, it is difficult to determine where students are in terms of skills and knowledge and to gauge the net effect of the NCLB. In addition, as noted by the Center on Education Policy (CEP), it is seemingly not possible to determine whether any gains observed in student achievement are directly attributable to the NCLB, as states and LEAs have implemented policies and programs of their own initiative and in response to the NCLB.⁵⁰ It is also not possible to determine what would have happened with respect to student achievement in the absence of the NCLB, as its provisions have probably affected every public school student in the nation.

A related issue is whether states have implemented challenging accountability systems. That is, questions have been raised about whether some states have implemented systems in which attaining proficiency is relatively easy, providing the appearance that the state is doing an excellent job educating students, while other states have implemented more rigorous accountability systems and results demonstrate that fewer students are succeeding academically. While the United States does not have a national test that examines what students know in a given subject area, all states that accept Title I-A funds are obligated to participate in the National Assessment of Educational Progress (NAEP), also referred to as the Nation's Report Card. While

⁵⁰ Center on Education Policy, *Has Student Achievement Increased Since 2002? State Test Score Trends Through 2006-07*, Washington, DC, February 2008, <http://www.cep-dc.org/document/docWindow.cfm?fuseaction=document.viewDocument&documentid=241&documentFormatId=4309>.

states have not written their content standards to be aligned with the NAEP standards, a comparison of NAEP test results for 4th and 8th grades in reading and mathematics has consistently shown that state accountability systems label substantially more students as being proficient than the results from the NAEP assessments would indicate. A recent study involving NAEP performance standards and state performance standards found that no state's reading performance standards were as high as the NAEP proficiency standards. In mathematics, Massachusetts (grades 4 and 8) and South Carolina (grade 8) were the only states where the state standards for proficiency were in the same range as the NAEP standards for proficiency.⁵¹

Researchers have developed a method for mapping each state's standard for proficiency in reading and mathematics onto the NAEP proficiency scale to permit a comparison of proficiency performance standards across states.⁵² Researchers found that the differences from state to state in the percentage of students scoring at the proficient level or higher could be explained by the variation in the degree of difficulty of performance standards for proficiency. States with higher performance standards based on the NAEP scale had fewer students scoring at the proficient level on the reading and mathematics assessments. This indicates that students of similar academic skill levels, but living in different states, are being evaluated against different performance standards. Similar results were found when another group of researchers compared state proficiency standards with benchmarks for two international assessments.⁵³ For example, researchers compared the performance standards in Massachusetts with those of the states with the lowest standards. They found that in the latter states, the 8th grade performance standards were equivalent to the 4th grade performance standards in Massachusetts.

Thus, it is clear that "proficient" does not carry the same meaning across states. In addition, in order to comply with the requirement that all students are proficient (however the state defines "proficient") in reading and mathematics by the end of the 2013-2014 school year, states can continue to exercise the flexibility they are granted by the NCLB to select performance standards that make it easier for more students to meet the proficiency bar.

Common Standards and Assessments

Concerns about the diversity of accountability systems across the nation have spurred a grassroots movement led by the National Governors Association (NGA) and the Council of Chief State School Officers (CCSSO) to develop common standards for reading and mathematics in grades K-12 (referred to as the common core standards) and upon which common assessments could be created to test student knowledge of the standards. Adoption and implementation of the standards and assessments is optional. Currently, 43 states and the District of Columbia have adopted the standards.⁵⁴ It remains to be seen how many of these states will ultimately implement the

⁵¹ U.S. Department of Education, National Center for Education Statistics, *Mapping State Proficiency Standards Onto NAEP Scales: 2005-2007*, NCES 2010-456, Washington, DC, October 2010, <http://nces.ed.gov/nationsreportcard/pdf/studies/2010456.pdf>.

⁵² U.S. Department of Education, National Center for Education Statistics, *Mapping State Proficiency Standards Onto NAEP Scales: 2005-2007*, NCES 2010-456, Washington, DC, October 2010, <http://nces.ed.gov/nationsreportcard/pdf/studies/2010456.pdf>.

⁵³ The two international tests used in the study were the Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS). For more information, see Gary W. Phillips, *International Benchmarking: State Education Performance Standards*, American Institutes for Research, Washington, DC, October 2010, http://www.air.org/files/AIR_Int_Benchmarking_State_Ed_Perf_Standards.pdf.

⁵⁴ The following states have not adopted the common standards: Alaska, Maine, Montana, Nebraska, North Dakota, (continued...)

standards and implement the standards as they were developed. In addition, 44 states and the District of Columbia have joined at least one of two groups currently working on developing common assessments for use by the 2014-2015 school year.⁵⁵ A decision to participate in the development process, however, does not necessarily translate into eventual adoption and use of the assessments.

The movement toward common standards and common assessments is not a federally led effort, *per se*. However, the movement clearly has the support of the Obama Administration. In its blueprint for the reauthorization of the ESEA,⁵⁶ the Administration proposes requiring states to adopt and implement common standards, presumably the aforementioned standards, or to have their standards vetted by a local university system. In addition, the Administration provided additional points to states competing for Race to the Top (RTTT) grants authorized by the ARRA if they adopted the common standards by a certain date. This incentive may have resulted in many more states agreeing to adopt the standards than would have otherwise occurred. Now that RTTT grants have been awarded to 11 states and the District of Columbia, however, it is possible that some states that failed to secure a grant may reconsider their adoption of the standards. With respect to support for common assessments, the Administration used a portion of the funds available for RTTT to run a \$350 million competitive grant program to support the development of common assessments based on common standards by non-federally affiliated groups.⁵⁷

Despite these grassroots efforts and actions by the Obama Administration to support the efforts, the end result will not yield a single set of national standards in reading and mathematics nor a single set of assessments in these subject areas. For example, states that adopt the common core standards are permitted to add additional standards of their own choosing to the common core standards.⁵⁸ Thus, each state adopting and implementing the common core standards could continue to have a unique set of state standards that share common elements with other adopting states. As a result of the RTTT common assessment competition, there will be at least two different assessments linked to the common core standards. It is unclear how the common assessments would accommodate any additional standards that states choose to add to the common core standards. Presumably, if states added additional standards to the common core standards, they would also want to determine how well students are mastering those standards. In addition, it is possible that multiple states could choose to use the same assessments to measure student performance but select different levels of performance on the assessments as indicating

(...continued)

Texas, and Virginia. Minnesota only adopted the common standards for reading/language arts. For more information, see <http://www.corestandards.org/in-the-states>.

⁵⁵ The following states did not join one of the two groups working on common assessments: Alaska, Minnesota, Nebraska, Texas, Virginia, and Wyoming. For more information, see U.S. Department of Education, "U.S. Secretary of Education Duncan Announces Winners of Competition to Improve Student Assessments," press release, September 2, 2010, <http://www.ed.gov/news/press-releases/us-secretary-education-duncan-announces-winners-competition-improve-student-asse>.

⁵⁶ U.S. Department of Education, *A Blueprint for Reform: Reauthorization of the Elementary and Secondary Education Act*, March 2010, <http://www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf>.

⁵⁷ For more information, see U.S. Department of Education, "U.S. Secretary of Education Duncan Announces Winners of Competition to Improve Student Assessments," press release, September 2, 2010, <http://www.ed.gov/news/press-releases/us-secretary-education-duncan-announces-winners-competition-improve-student-asse>.

⁵⁸ States adopting the common core standards are required to adopt all of the standards but may add an additional 15% of its own materials to the standards. For more information, see PowerPoint presentation dated March 2010, available online at <http://www.corestandards.org/about-the-standards>.

proficiency (e.g., one state could say that students must get 75% of the questions correct on the common assessment to be considered proficient, while another state could set the bar at 50%).

It is important to note that neither the common core standards movement nor the assessments movement is proposing a common curriculum. Decisions regarding how standards are taught to students and how students are prepared for assessments would remain a state and local decision. However, if enough states that adopted the common core standards actually implement the standards, it is possible that states could work together to develop, or textbook publishers and other organizations that develop materials for classroom use may develop, materials that are clearly aligned with the common core standards; thus, these entities would possibly contribute to a de facto national curriculum.

In the long term, however, if the push for more commonality gains momentum, there are several issues that remain unresolved regarding the common core standards. For example, who is responsible for maintaining and updating the standards? Who pays for this? How often should changes be made to the standards, as changes to the standards would require concomitant changes to assessments, teacher training, and curricula? If the common core standards are included in the ESEA, would there be a federal role in maintaining or updating the standards? Similar questions may apply to any common assessments that are developed, particularly with respect to the responsibility for maintaining and updating the assessments if the standards to which they are aligned change.

If ESEA reauthorization includes common standards or common assessments as part of a revised accountability system, this may increase the number of states that actually implement the standards and ultimately use the assessments. This may be particularly appealing if the adoption of common standards and assessments involves some sort of tradeoff, such as increasing state and local flexibility with respect to the consequences applied to schools and LEAs that fail to meet AYP (or some other required performance target). Alternatively, Congress could choose not to incorporate common standards or common assessments into a revised ESEA, leaving the common standards and assessments movements at the grassroots level. Under this scenario, Congress could allow the common standards and assessments to be used to meet whatever accountability requirements continue to apply without actually addressing the issues of the common standards or assessments in statutory language.

Absolute Versus Differentiated Consequences

Under the NCLB, LEAs and states simply do or do not meet AYP standards. There is generally no distinction between those that fail to meet only one or two required performance or participation thresholds to a marginal degree versus those that fail to meet numerous thresholds to a substantial extent. Several analysts have suggested that a more nuanced grading scale be allowed (e.g., a division of schools or LEAs failing to make AYP into higher versus lower priority categories, or grades ranging from A to F), as is used in several state accountability systems. A major complication associated with such an approach is determining at what point on such a scale the current automatic consequences (e.g., school choice or SES) are invoked. An additional consideration under the NCLB's current performance-based accountability system is that the same set of consequences applies to all schools regardless of whether they failed to make AYP for one subgroup or all students. That is, there is generally no distinction between those that fail to meet only one or two required performance thresholds to a marginal degree versus those that fail to meet numerous thresholds to a substantial extent.

Further, studies have determined that when otherwise-similar public schools are compared, those with a wider variety of demographic groups are substantially less likely to meet AYP standards.⁵⁹ However, without specific requirements for achievement gains by each of the major student groups, it is possible that insufficient attention would be paid to the performance of the disadvantaged student groups among whom improvements are most needed, and for whose benefit the Title I-A program was established.

As more schools fail to make AYP and are identified for improvement, the funds that LEAs must reserve for public school choice and SES may not be sufficient to help all eligible students. In addition, the public school transfer option may become a moot point, as few LEAs may have enough schools making AYP into which students can transfer. Two studies examined the availability of school choice options during the 2004-2005 school year. The first study⁶⁰ found that 39% of LEAs that were required to offer choice failed to do so. Among LEAs required to offer choice, 20% reported that they had no non-identified schools within the LEA to which students could transfer either because they only had one school per grade level or all schools in the LEA had been identified for improvement. Other LEAs noted that a lack of space in non-identified schools or an inability to arrange inter-district transfers were major challenges to implementing the choice provision. A second study⁶¹ echoed these findings and noted that finding physical space and adhering to class size requirements in non-identified schools made it difficult to comply with the choice requirements. Given that more schools have been identified for improvement since the 2004-2005 school year, it is likely that it has only become more difficult for LEAs to comply with the public school choice requirements included in Title I-A.

A related issue is the capacity of states and LEAs to provide assistance to schools and LEAs that fail to make AYP for two consecutive years or more. As more schools and LEAs are subject to performance-based accountability consequences, the ability of states and LEAs to provide assistance may be diminished. It has been suggested that the consequences need to be better differentiated based on the reasons for failure and that assistance may need to be focused on the lowest-performing schools.

Finally, the focus of the current performance-based accountability system has been on consequences for failing to meet required levels of performance. While performance-based rewards are authorized under the NCLB, they are apparently little used, and the current focus is on a variety of performance-based sanctions. The Obama Administration's blueprint for reauthorization would place a greater emphasis on incentivizing improved performance through the use of rewards. Congress could choose to alter its approach to obtaining compliance with ESEA requirements to focus on positive incentives rather than performance-based sanctions or, perhaps, it could achieve a more equitable balance between the two approaches.

⁵⁹ See, for example, U.S. Department of Education, *State and Local Implementation of the No Child Left Behind Act Volume IX—Accountability under NCLB: Final Report*, 2010, pp. 55-57, <http://www2.ed.gov/rschstat/eval/disadv/nclb-accountability/nclb-accountability-final.pdf>.

⁶⁰ U.S. Department of Education, Institute of Education Sciences, *National Assessment of Title I: Interim Report, Volume I: Implementation*, February 2003, p. 65, <http://www2.ed.gov/rschstat/eval/disadv/title1interimreport/vol1.pdf>.

⁶¹ Center on Education Policy, *From the Capital to the Classroom: Year 3 of the No Child Left Behind Act*, 2005, http://www.cep-dc.org/_data/n_0001/resources/live/Supplemental%20Education%20Services%20and%20Choice.pdf.

Proficiency for All Students

Without an ultimate goal of having all students reach a proficient or higher level of achievement by a specific date, states might establish relative goals that require little or no net improvement over time. A demanding goal might maximize efforts toward improvement by state public school systems, even if the goal is not met. Nevertheless, a goal of having all students at a proficient or higher level of achievement, within any specified period of time, may be criticized as being unrealistic, if one assumes that proficiency has been set at a challenging level. It is very likely that many states, schools, and LEAs will not meet the NCLB's 2014 AYP goal, unless state standards of proficient performance are significantly lowered or states are allowed by ED to aggressively pursue the use of statistical techniques such as setting high minimum group sizes and confidence intervals.

The assessment data that have been collected since the enactment of the NCLB could possibly now be used to develop new goals for student performance that are grounded in empirical evidence of what levels of growth may be achievable. For example, in reauthorizing the ESEA, Congress could rely on empirically developed thresholds for student performance. Alternatively, required progress could be benchmarked against a school or LEA that is already considered to be effective. Regardless of the strategy used to develop academic achievement goals for students, the strategy may need to recognize that as the percentage of students who are required to be proficient increases, it will be harder to realize large gains in the percentage of students who are proficient.

Incentives to Focus on Proficiency

A potentially positive outcome of the NCLB performance-based accountability system has been an increased focus on the proficiency of *all students*, including previously underserved groups such as disadvantaged students, students with disabilities, and LEP students. By disaggregating assessment results by subgroup, the achievement of various student groups is transparent and schools are held accountable for the achievement of these previously underserved groups.

Although schools and LEAs are held accountable for the achievement of students in various student groups, they are not held accountable for students at all levels of achievement. Under the NCLB performance-based accountability system, states set targets for the percentage of students achieving proficiency in reading and mathematics. States are not currently required to set targets for the percentage of students achieving other levels of achievement, such as "basic" or "advanced." The accountability system, therefore, does not reward schools or LEAs for increasing the percentage of students moving from achievement levels that are far below basic to the "basic" level of achievement. Likewise, the system does not reward schools or LEAs for increasing the percentage of students moving from "proficient" to "advanced" levels of achievement.

Because the goal of the current system is for 100% of students to become proficient by school year 2013-2014, schools and teachers may target instructional time and resources toward students who are nearing proficiency rather than distributing resources equally across students at all achievement levels. Congress could choose to alter the requirements of the current performance-based accountability system to provide a more balanced approach in rewarding gains in student achievement. For example, a performance-based accountability system could be designed to reward schools for increasing student growth in achievement, even if the ultimate level of

achievement does not reach the cut-off score for proficiency. Schools and LEAs could similarly be rewarded for increasing student growth above the level of proficiency.

There has been an increasing interest in implementing state accountability systems that incorporate measures of student growth rather than one end-of-year assessment that determines proficiency. Currently, a limited number of states use growth models for accountability purposes;⁶² however, ED's requirements for growth models are considered to be relatively restrictive.⁶³ The current Administration has also expressed an interest in assessing student growth rather than static achievement.⁶⁴ To date, the current Administration has not provided details on how new measures of student growth would differ from the previous Administration's growth model requirements. As such, it is difficult to gauge how this renewed interest in growth models would change current policy. If Congress chooses to expand the use of growth models in state accountability systems, the assessment requirements for the growth model must be considered. Some growth models require assessing students more than once per year; some growth models require assessments to be aligned across multiple grade levels. If states are granted the flexibility to use less restrictive growth models, Congress may want to consider implementing a review procedure that ensures that the states have the appropriate assessments in place to inform their growth model.

Assessment and the Narrowing of the Curricular Focus⁶⁵

The NCLB placed student assessment at the heart of federal education policy. One potentially positive outcome of using student assessments has been an increased focus on state content standards and teaching to those standards. There is some evidence that using content standards and assessments can help teachers focus their instruction and obtain feedback on the effectiveness of instruction.⁶⁶ On the other hand, the use of assessments in test-based accountability systems may be affecting the curriculum in less desirable ways. Because assessments are aligned with state content standards, there may be a risk that "teaching to the standards" becomes "teaching to the test." The practice of "teaching to the test"—whether intentional or unintentional—may narrow the curriculum in several ways. For example, teachers may reallocate instructional time towards tested subjects and away from non-tested subjects. Surveys of teachers have consistently reported that their instruction emphasizes reading and mathematics over other subjects like history, foreign language, and arts.⁶⁷ "Teaching to the test" can also lead to a narrowing of the

⁶² States with currently approved growth models include Alaska, Arizona, Arkansas, Colorado, Delaware, Florida, Iowa, Michigan, Minnesota, Missouri, North Carolina, Ohio, Pennsylvania, Tennessee, and Texas. For more information, see <http://www2.ed.gov/news/pressreleases/2009/01/01082009a.html>.

⁶³ The growth model requirements were outlined by Secretary Spellings under the previous Administration.

⁶⁴ U.S. Department of Education, *A Blueprint for Reform: Reauthorization of the Elementary and Secondary Education Act*, March 2010, <http://www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf>.

⁶⁵ For general information on assessment in elementary and secondary education, see CRS Report R40514, *Assessment in Elementary and Secondary Education: A Primer*, by (name redacted).

⁶⁶ L. Mabry, J. Poole, and L. Redmond, et al., "Local Impact of State Testing in Southwest Washington," *Education Policy Analysis Archives*, vol. 11, no. 22 (July 18, 2003), <http://epaa.asu.edu/epaa/v11n22/>.

⁶⁷ Laura Hamilton, "Assessment as a Policy Tool," *Review of Research in Education*, vol. 27 (2003), pp. 25-68; Laura S. Hamilton and Mark Berends, *Instructional Practices Related to Standards and Assessment*, RAND, WR-374-EDU, Washington, DC, April 2006, http://www.rand.org/pubs/working_papers/2006/RAND_WR374.pdf; Patricia Velde Pederson, "What is Measured Is Treasured: The Impact of the No Child Left Behind Act on Nonassessed Subjects," *Clearing House: A Journal of Educational Strategies, Issues and Ideas*, vol. 80, no. 6 (July/August 2007), pp. 287-291; Jennifer McMurrer, *Choices, Changes, and Challenges: Curriculum and Instruction in the NCLB Era*, Center on (continued...)

curriculum within tested subjects. If teachers are aware that the assessments measure certain skills *within* the reading and mathematics standards, it is more likely that these skills would be taught—possibly at the expense of other, non-tested skills that are more difficult to measure. For example, if teachers know that a mathematics assessment typically assesses basic facts, it is less likely that higher-order skills, such as problem solving, would be adequately emphasized in the curriculum. A typical multiple-choice test that consists of 40 to 50 items per content area translates into roughly one or two test items per standard, leaving many curricular objectives within the standard untested—particularly those that are difficult to measure.⁶⁸

A related concern is the potential for score inflation. Score inflation is a phenomenon in which scores on high-stakes assessments, such as state assessments used in accountability systems, increase at faster rates than scores on low-stakes assessments.⁶⁹ If teachers are, in fact, “teaching to the test” (i.e., the state assessment), it is likely that scores will increase faster on the state assessment than on other tests of student achievement. Studying the prevalence of score inflation is difficult because schools and LEAs may be reluctant to give researchers access to test scores for the purpose of investigating possible inflation. Nevertheless, several studies have documented the problem of score inflation by comparing gains on state assessments (high stakes) to those made on the NAEP (low stakes).⁷⁰ Studies have consistently reported discrepancies in the overall level of student achievement, the size of student achievement gains, and the size of the achievement gap. The discrepancies indicate that student scores on state assessments may be inflated and that these inflated scores may not represent true achievement gains.⁷¹

One possible way to reduce the problem of score inflation is to consistently use a low-stakes audit assessment, such as NAEP, to corroborate gains on state assessments.⁷² If gains on state

(...continued)

Education Policy, December 2007, http://www.cep-dc.org/_data/n_0001/resources/live/07107%20Curriculum-WEB%20FINAL%207%2031%2007.pdf.

⁶⁸ Henry May, Irma Perez-Johnson, and Joshua Haimson, et al., *Using State Tests in Education Experiments: A Discussion of the Issues*, U.S. Department of Education, NCEE 2009-013, Washington, DC, November 2009, pp. A-8, <http://ies.ed.gov/ncee/pdf/2009013.pdf>.

⁶⁹ An assessment is considered to be high stakes if there are consequences for individuals, schools, LEAs, or states based on the outcome of the assessment. State assessments for AYP and state exit exams are examples of high-stakes assessments. An assessment is considered to be low-stakes if there are no consequences for the outcome of the assessment. NAEP and various international assessments are examples of low-stakes assessments.

⁷⁰ NAEP is administered in grades 4, 8, and 12. Studies comparing student performance on state assessments to the NAEP make comparisons only at grades 4 and 8. B. Fuller, K. Gesicki, and E. Kang, et al., *Is the No Child Left Behind Act Working? The Reliability of How States Track Achievement*, University of California, Berkeley PACE, Working Paper 06-1, Berkeley, CA, 2006; S.P. Klein, Linda S. Hamilton, and Daniel F. McCaffrey, et al., *What Do Test Scores in Texas Tell Us?*, RAND, Santa Monica, CA, 2000; Daniel Koretz and S. I. Barron, *The Validity of Gains on the Kentucky Instructional Results Information System (KIRIS)*, RAND, MR-792-PCT/FF, Santa Monica, CA, 1998; Robert L. Linn and C. Haug, “Stability of School-building Accountability Scores and Gains,” *Educational Evaluation and Policy Analysis*, vol. 24, no. 1 (2002), pp. 29-36.

⁷¹ Critics of the audit approach may argue that NAEP and state assessments measure reading and mathematics much differently, and a gain on one assessment would not necessarily be reflected on the other assessment. That is, the two types of assessments measure achievement against different content and performance standards, and a gain on the state assessment may reflect the specific content that is emphasized by the state standards. Others, however, may argue that it is unlikely that the assessments are different to the extent that a gain on one assessment would not likely coincide with a gain on the other assessment. That is, it is unlikely that true achievement gains would be demonstrated on only one assessment.

⁷² Daniel Koretz, *Measuring Up: What Educational Testing Really Tells Us* (Cambridge, MA: Harvard University Press, 2008), pp. 247-248.

assessments generalize to another audit assessment, it increases the likelihood that gains are due to true achievement gains. This type of corroboration may help policymakers separate the policies that lead to true student achievement from those that lead to score inflation.

It can be argued that the assessment requirements of the NCLB accountability system have become burdensome and that more assessment decisions should be made at the state and local levels. Alternatively, it can be argued that the assessment requirements of the NCLB have the potential to drive curriculum, and that Congress should consider expanding the requirements to include subjects like writing, history, civics, and the arts. If the requirements are expanded to include additional subjects, these requirements may provide an incentive to states and LEAs to broaden school curricula. Even if the federal assessment requirements were expanded to include additional subjects, however, it is unclear whether scores from these new assessments would be included in the NCLB accountability system. For example, states are currently required to assess science, but the science score is not counted in AYP determinations. If scores from additional assessments are not counted in AYP determinations, it is unclear whether an assessment without consequences would broaden the curriculum.

It can be argued that the current assessments should be redesigned to measure more higher-order thinking skills. States currently administer standardized, multiple-choice assessments, largely due to the ease of administration, scoring, and reporting; however, this practice is sometimes considered undesirable because it may not adequately measure important higher-order thinking and skills. As such, supporters of innovative assessments may argue that the federal government should support other types of assessments that are more likely to assess higher-order thinking skills, such as extended constructed response and performance-based assessments.⁷³ Others may argue that the federal government should require the use of multiple measures.⁷⁴ That is, in addition to a state assessment score, AYP determinations could include data from a portfolio of student work, a performance assessment, or observations of student behavior and achievement in the classroom.

Technical Quality of Assessments⁷⁵

Using student assessments in performance-based accountability systems provides a method of determining whether students are meeting certain academic targets. It highlights areas of excellence and areas of need within schools, LEAs, and states. Without the use of assessments, states would not have a baseline measurement of academic achievement to gauge whether certain instructional, curricular, or policy changes led to improved student outcomes. Nevertheless, the use of assessments in performance-based accountability systems is not without limitations. For example, most states currently use one, end-of-year assessment score in reading and mathematics

⁷³ See, for example, George H. Wood, Linda Darling-Hammond, Monty Neil, et al., *Refocusing Accountability: Using Local Performance Assessments to Enhance Teaching and Learning for Higher Order Skills*, The Forum for Education and Democracy, May 2007, <http://www.forumforeducation.org/node/368>. An assessment with *extended constructed response* may require a student to write and essay or design an experiment. A *performance-based assessment* may require a student to demonstrate skills and knowledge in an “on-demand” setting, such as a debate or a laboratory experiment. There are many different examples of extended constructed response and performance-based assessments that are not mentioned herein.

⁷⁴ Congressman Miller and Congressman McKeon proposed the use of multiple indicators or multiple measures in the 2007 discussion draft to reauthorize the NCLB.

⁷⁵ For a guide to the technical validity of educational assessments, see AERA, APA, and NCME, *Standards for Educational and Psychological Testing* (Washington, DC: American Psychological Association, 1999).

to inform the accountability system.⁷⁶ Because of the emphasis placed on a single assessment score, the appropriate interpretation of that score is critical. If assessments are not used and interpreted accurately, changes in student achievement scores may be used inappropriately to support or dismiss certain school reform efforts.

Given the emphasis placed on the results of these assessments, it is important that teachers, administrators, and policymakers understand the limitations of the assessment. Educational assessments can be limited by issues of technical quality, such as validity, reliability, and fairness.

Validity. Validity is arguably the most important concept to understand when evaluating educational assessment; it is not a property of the assessment itself, but rather a property of the interpretation of the result of the assessment. Validity refers to the degree to which a certain interpretation of an assessment score is appropriate and meaningful. That is, from a single assessment score, one could make a valid or an invalid interpretation. For example, state assessment results from a school are a representation of what students in the school know and can do relative to the state academic content standards. Suppose a group of students in School A scored well on these assessments and a group of students in School B scored poorly. One possible valid inference from this result is that students from School A were more likely to demonstrate proficiency in the academic content than students from School B. There are, however, many possible inferences that may be less valid. For example, one could infer that School A had a better academic curriculum than School B. Or, one could infer that School A had better teachers than School B. Neither of these inferences may be valid because state assessments were designed for the purpose of determining students' mastery of the content standards of the state, not for the purposes of evaluating teachers or curriculum. The validity of an inference, therefore, is tied inextricably to the purpose for which the test was created.

Reliability. Reliability refers to the consistency of measurement of an assessment. It describes the precision with which assessment results are reported and is a measure of the certainty that the results are accurate. The concept of reliability presumes that each student has a *true score* for any given assessment. The score on the assessment, or the *observed score*, is simply an approximation of a student's true score. The higher the reliability of the assessment, the more likely it is that a student's observed score reflects the true score. That is, the higher the reliability of the assessment, the more likely it is that an assessment score reflects student achievement.

Fairness. Fairness is a term that has no technical meaning in assessment procedures, and the concept has multiple interpretations. In general, fairness refers to the concept of equal treatment for all students who participate in an educational assessment. For example, the assessment process may lack fairness if it is biased against certain individuals. In a reading test, unnecessarily complex vocabulary can potentially bias a test against LEP students (e.g., using the word "parakeet" where "bird" would suffice). Or, the assessment process may lack fairness if certain individuals have been provided more test preparation opportunities. Some schools may emphasize test-taking skills and expose the students to items from previous state assessments to familiarize them with the general form and content. Other schools, however, may not emphasize test-taking skills, opting instead to use the time to deliver more instruction. The extent to which test preparation boosts scores on state assessments is unknown; however, if one school conducted

⁷⁶ States are required to use at least one other academic indicator in their accountability system (e.g., high school graduation rate, attendance, etc.). Some states use growth models in their accountability system. In some growth models, students are tested twice per year to determine growth. In other models, students are tested once per year and the prior year's score is used with the current year's score to determine growth.

extensive test preparation and another did not, comparisons between the two schools may not be “fair.”⁷⁷

The results of state assessments are a major component of the NCLB performance-based accountability system. Consequences for schools and LEAs are based primarily on the results of state assessments. As such, test users have a responsibility to examine the validity, reliability, and fairness of an assessment to draw defensible conclusions about student achievement. If a subgroup of students does not make AYP based on an assessment that lacked adequate evidence of reliability or fairness, schools and LEAs would still be subjected to the increasingly severe consequences of the performance-based accountability system. If the consequences applied to schools and LEAs are based on assessment scores that are invalid, it can be argued that current school reform efforts are not targeted appropriately. Some schools or LEAs that provide relatively sound educational opportunities for students may be required to implement a series of reform efforts because the assessments used within the accountability system did not accurately and fairly measure the achievement of all students, while other schools and LEAs may fail to be identified as needing to improve their educational offerings.

Teacher Evaluation and Accountability⁷⁸

One of the goals of the NCLB is to improve educational outcomes for all students. Policymaking at the federal level reflects a growing awareness that improving educational outcomes depends greatly upon increasing the quality of classroom instruction. In establishing the student performance standards and accountability provisions in the NCLB, Congress recognized that the success of these reforms rests largely on teachers’ knowledge and skills. Thus, in enacting the NCLB, Congress amended the ESEA to establish a requirement that all teachers be *highly qualified*. Furthermore, Congress required that states establish a plan for the equitable distribution of highly qualified teachers across classrooms and schools. The equitable distribution of highly qualified teachers was intended, “to ensure that poor and minority children are not taught at higher rates than other children by inexperienced, unqualified, or out-of-field teachers.”⁷⁹

The NCLB highly qualified teacher requirement has been the cornerstone of federal teacher policy for nearly a decade. In that time, the requirement has come to be seen by many as a minimum standard for entry into the profession (rather than a goal to which teachers might aspire) and a growing body of research has revealed its underlying emphasis on teachers’ credentials to be weakly correlated with student achievement.

Meanwhile, federal interest has begun to shift from a focus on teacher input (i.e., quality) to teacher output (i.e., effectiveness). The most recent congressional action in this area came with the passage of the American Recovery and Reinvestment Act and, in particular, enactment of the RTTT program. Within the RTTT program, ED provided the first federal definition of “effective teacher” and required applicants of this program to measure a teacher’s effectiveness based on

⁷⁷ For more information on fairness in assessment, see footnote 65.

⁷⁸ This section of the report draws on work done by Jeff Kuenzi, Specialist in Education Policy. For more information on teacher policy, see CRS Report RL33333, *A Highly Qualified Teacher in Every Classroom: Implementation of the No Child Left Behind Act and Reauthorization Issues for the 112th Congress*, by (name redacted); and CRS Report R41267, *Elementary and Secondary School Teachers: Policy Context, Federal Programs, and ESEA Reauthorization Issues*, by (name redacted).

⁷⁹ ESEA, Section 1111(b)(8)(C).

student growth in achievement. In addition, applicants of this program were to use the measurement of teacher effectiveness within an overall teacher evaluation system that would inform rewards for and recruitment, development, and retention of teachers.

ED provided RTTT applicants with an incentive to become more systematic about using student data to inform teacher instruction and to measure teacher effectiveness. One specific method that is used by some school districts (and at least one state)⁸⁰ to measure teacher effectiveness based on student achievement data is value-added modeling (VAM). VAM is a set of statistical approaches that seek to isolate a teacher's effect on student achievement. This method of modeling is seen as promising because it has the potential to promote education reform and to create a more equitable accountability system that holds teachers and schools accountable for the aspects of student learning that are attributable to effective teaching while not holding teachers and schools accountable for factors outside of their control (e.g., the potential impact of socioeconomic status on student learning). Although there are other methods for assessing the effectiveness of teachers, VAM has garnered increasing attention in education research and policy, and many states proposed VAM systems in RTTT applications.⁸¹

To date, the shift from input-based accountability (i.e., teacher quality) to output-based accountability (i.e., teacher effectiveness) has been an initiative led by the Administration. Congress has not enacted similar provisions, although this may receive attention during ESEA reauthorization. During reauthorization, Congress may choose to maintain the highly qualified teacher requirements as they are, replace or supplement these requirements, or include these requirements as part of a new, more comprehensive set of quality/performance requirements. In addition, Congress may revisit the requirement for the equitable distribution of highly qualified teachers to ensure that it is functioning as intended to ensure that poor and minority children are not taught by inexperienced or unqualified teachers at higher rates than other children.⁸²

On the other hand, if Congress chooses to move toward outcome-based accountability, it may consider whether to codify ED's definition of "effective teacher" or whether to consider alternate definitions. If ED's definition is adopted, reauthorization legislation has the opportunity to clarify the weight that student growth in achievement should have in the definition of an effective teacher. Furthermore, if Congress chooses to adopt any definition of an effective teacher, the equitable distribution of "highly qualified teachers" may need to be revised to reflect the change to an emphasis on "effective teachers."

Finally, Congress may consider whether the federal government will have a sustained role in incentivizing states to develop new teacher evaluation systems that would inform rewards for and recruitment, development, and retention of teachers. Through the Teacher Incentive Fund (TIF), Congress currently provides grants to support compensation reforms for teachers.⁸³ TIF grants

⁸⁰ Several large school districts – including Chicago, Denver, and the District of Columbia – use value-added systems for teacher evaluations. Tennessee uses a statewide value-added system for teacher evaluations.

⁸¹ For more information on VAM, see CRS Report R41051, *Value-Added Modeling for Teacher Effectiveness*, by (name redacted) and (name redacted).

⁸² Although some claim that high-poverty schools are less likely to have highly qualified teachers, given that nearly all teachers are considered highly qualified, it is not clear whether this requirement has proven useful. According to the most recent figures, from the 2007-2008 school year, 95% of all core academic classes are taught by highly qualified teachers; the figure is 99% in more than a dozen states. U.S. Department of Education, *A Summary of Highly Qualified Teacher Data*, Washington, DC, May 2009, <http://www2.ed.gov/programs/teacherqual/data2009.doc>.

⁸³ TIF was authorized by P.L. 109-149.

must consider gains in student achievement, classroom evaluations conducted multiple times during each school year, and other factors that provide educators with incentives to take on additional responsibilities and leadership roles. Through RTTT, ED required that grantees of this program go beyond TIF requirements and use evaluations to inform high-stakes decisions such as granting tenure, awarding full certification, and removing ineffective teachers. In practice, some school districts are using teacher evaluation systems based on VAM for these high-stakes decisions. Some have cautioned against using VAM for tenure, certification, and removal decisions because the measurement of teacher effectiveness may not be precise enough to support these decisions without additional corroborating evidence.⁸⁴ If the measurement of teacher effectiveness lacks precision, Congress may consider the appropriate weight given to teacher effectiveness in an overall teacher accountability system.

⁸⁴ For more information on VAM, see CRS Report R41051, *Value-Added Modeling for Teacher Effectiveness*, by (name redacted) and (name redacted); and Steven Glazerman, Susanna Loeb, and Dan Goldhaber, et al., *Evaluating Teachers: The Important Role of Value-Added*, Brown Center on Education Policy at Brookings, Washington, DC, November 17, 2010, http://www.brookings.edu/~media/Files/rc/reports/2010/1117_evaluating_teachers/1117_evaluating_teachers.pdf.

Appendix. Related Tables

**Table A-1. Percent of All Public Schools and LEAs that Did Not Make AYP:
2008-2009 School Year**

State	Percentage of All Public Schools That Did Not Make AYP	Percentage of All LEAs That Did Not Make AYP
Alabama	13.3%	2.3%
Alaska	43.8%	—
Arizona	25.6%	39.2%
Arkansas	45.7%	9.6%
California	49.9%	64.0%
Colorado	43.9%	53.8%
Connecticut	41.1%	32.0%
Delaware	33.8%	36.8%
District of Columbia	74.7%	85.7%
Florida	76.6%	—
Georgia	14.0%	60.2%
Hawaii	64.4%	—
Idaho	33.7%	55.0%
Illinois	40.8%	49.4%
Indiana	49.9%	16.8%
Iowa	30.2%	9.7%
Kansas	12.2%	11.5%
Kentucky	38.4%	57.5%
Louisiana	9.2%	56.1%
Maine	35.3%	0.9%
Maryland	23.0%	84.0%
Massachusetts	62.2%	79.0%
Michigan	9.2%	36.3%
Minnesota	53.7%	61.1%
Mississippi	35.4%	88.8%
Missouri	62.6%	72.9%
Montana	26.4%	32.1%
Nebraska	11.7%	19.3%
Nevada	42.8%	35.3%
New Hampshire	54.0%	48.2%
New Jersey	35.1%	16.5%
New Mexico	68.2%	52.7%

State	Percentage of All Public Schools That Did Not Make AYP	Percentage of All LEAs That Did Not Make AYP
New York	11.7%	4.9%
North Carolina	28.9%	89.6%
North Dakota	25.2%	23.0%
Ohio	39.4%	51.8%
Oklahoma	10.6%	11.6%
Oregon	29.9%	52.6%
Pennsylvania	21.5%	7.2%
Puerto Rico	93.5%	—
Rhode Island	18.9%	22.5%
South Carolina	49.8%	—
South Dakota	20.7%	5.7%
Tennessee	20.3%	16.9%
Texas	5.0%	17.3%
Utah	16.8%	13.1%
Vermont	29.0%	26.9%
Virginia	28.1%	54.9%
Washington	58.2%	69.3%
West Virginia	19.7%	94.5%
Wisconsin	6.7%	0.9%
Wyoming	27.4%	8.3%
National average	33.7%	36.3%

Source: Table prepared by CRS, November 24, 2010, based on data available from the U.S. Department of Education, Education Data Express System. The original source of the data is the EDFacts/Consolidated State Performance Report, 2008-09, available online at <http://www2.ed.gov/admins/lead/account/consolidated/index.html>.

Notes: Due to missing values, national-level data may not accurately represent national trends. States with missing data are not included in the national average calculation.

A dash (—) indicates that the data are not available.

**Table A-2. Number of Title I-A Schools in Improvement, Corrective Action, or Restructuring During the 2009-2010 School Year
Based on Adequate Yearly Progress Determinations from the 2008-2009 School Year**

State	Number of Title I-A Schools: 2008-2009 School Year	Number of Title I-A Schools in Improvement, Year 1	Number of Title I-A Schools in Improvement, Year 2	Number of Title I-A Schools in Corrective Action	Number of Title I-A Schools Planning for Restructuring	Number of Title I-A Schools Implementing Restructuring	Total Number of Title I-A Schools in Restructuring	Total Number of Title I-A Schools in Need of Improvement	Percentage of Title I-A Schools in Need of Improvement
Alabama	855	9	5	9	11	4	15	38	4.4%
Alaska	282	26	8	10	13	61	74	118	41.8%
Arizona	1,176	49	66	55	36	40	76	246	20.9%
Arkansas	812	101	86	66	50	100	150	404	49.8%
California	6,013	748	309	323	328	1,075	1,403	2,783	46.3%
Colorado	603	81	17	21	8	37	45	164	27.2%
Connecticut	507	64	48	31	14	80	94	237	46.7%
Delaware	106	4	6	1	-	3	3	14	13.2%
District of Columbia	189	19	26	25	22	37	59	129	68.3%
Florida	1,361	120	97	96	145	541	686	999	73.4%
Georgia	1,292	-	-	-	-	-	-	-	-
Hawaii	180	16	3	13	8	65	73	107	59.4%
Idaho	375	55	39	35	27	8	35	164	43.7%
Illinois	2,272	209	92	48	36	336	372	721	31.7%
Indiana	843	-	-	-	-	-	-	258	30.6%
Iowa	648	100	7	5	7	1	8	120	18.5%
Kansas	659	11	10	4	3	4	7	32	4.9%
Kentucky	820	29	21	19	7	30	37	106	12.9%
Louisiana	1,059	32	26	1	1	15	16	75	7.1%
Maine	421	37	9	4	1	4	5	55	13.1%
Maryland	358	10	10	7	6	38	44	71	19.8%

State	Number of Title I-A Schools: 2008-2009 School Year	Number of Title I-A Schools in Improvement, Year 1	Number of Title I-A Schools in Improvement, Year 2	Number of Title I-A Schools in Corrective Action	Number of Title I-A Schools Planning for Restructuring	Number of Title I-A Schools Implementing Restructuring	Total Number of Title I-A Schools in Restructuring	Total Number of Title I-A Schools in Need of Improvement	Percentage of Title I-A Schools in Need of Improvement
Massachusetts	985	146	111	102	100	188	288	647	65.7%
Michigan	1,922	-	106	18	-	84	84	208	10.8%
Minnesota	862	132	83	36	19	13	32	283	32.8%
Mississippi	707	37	25	5	4	3	7	74	10.5%
Missouri	1,125	212	101	65	24	56	80	458	40.7%
Montana	630	71	13	7	4	40	44	135	21.4%
Nebraska	466	14	1	1	-	-	-	16	3.4%
Nevada	154	32	15	23	41	20	61	131	85.1%
New Hampshire	238	57	32	24	17	2	19	132	55.5%
New Jersey	1,370	130	69	37	31	73	104	340	24.8%
New Mexico	570	93	60	54	69	118	187	394	69.1%
New York	3,262	100	69	47	36	175	211	427	13.1%
North Carolina	1,133	160	149	110	54	48	102	521	46.0%
North Dakota	305	5	7	2	-	14	14	28	9.2%
Ohio	2,133	186	176	191	99	123	222	775	36.3%
Oklahoma	1,122	20	5	3	5	2	7	35	3.1%
Oregon	574	47	8	-	7	10	17	72	12.5%
Pennsylvania	1,821	122	64	47	28	130	158	391	21.5%
Puerto Rico	1,487	224	65	86	137	430	567	942	63.3%
Rhode Island	154	13	6	5	8	12	20	44	28.6%
South Carolina	491	67	51	29	30	88	118	265	54.0%
South Dakota	344	16	4	11	2	21	23	54	15.7%

State	Number of Title I-A Schools: 2008-2009 School Year	Number of Title I-A Schools in Improvement, Year 1	Number of Title I-A Schools in Improvement, Year 2	Number of Title I-A Schools in Corrective Action	Number of Title I-A Schools Planning for Restructuring	Number of Title I-A Schools Implementing Restructuring	Total Number of Title I-A Schools in Restructuring	Total Number of Title I-A Schools in Need of Improvement	Percentage of Title I-A Schools in Need of Improvement
Tennessee	982	57	22	8	9	11	20	107	10.9%
Texas	5,093	159	66	58	39	30	69	352	6.9%
Utah	246	9	3	-	-	-	-	12	4.9%
Vermont	223	35	4	15	2	7	9	63	28.3%
Virginia	711	58	26	6	7	6	13	103	14.5%
Washington	923	237	142	14	35	40	75	468	50.7%
West Virginia	356	6	6	1	3	7	10	23	6.5%
Wisconsin	1,152	41	20	9	4	5	9	79	6.9%
Wyoming	179	15	2	1	3	1	4	22	12.3%
Total	52,551	4,221	2,396	1,788	1,540	4,236	5,776	14,561	27.7%

Source: Table prepared by CRS, November 22, 2010, based on data available from the U.S. Department of Education, Education Data Express System. The original source of the data is the EDFacts/Consolidated State Performance Report, 2008-09, available online at <http://www2.ed.gov/admins/lead/account/consolidated/index.html>.

Notes: A Title I-A school is a school that received federal funds through the Title I-A program authorized by the Elementary and Secondary Education Act (ESEA). States are only required to apply the performance-based outcome accountability requirements to Title I-A schools that fail to make adequate yearly progress. There are no data for Indiana because the state developed new improvement statuses under a differentiated accountability pilot that do not correspond with the existing improvement statuses. Due to missing values, national-level data may not accurately represent national trends.

Author Contact Information

(name redacted)
Specialist in Education Policy
[redacted]@crs.loc.gov, 7-....

Acknowledgments

The authors wish to acknowledge previous work conducted by (name redacted) and James B. Stedman, former Specialists in Education Policy at CRS, and Jeff Kuenzi, Specialist in Education Policy, on issues related to those included in this report. Erin Lomax, former CRS Specialist in Education Policy, also contributed to this report.

EveryCRSReport.com

The Congressional Research Service (CRS) is a federal legislative branch agency, housed inside the Library of Congress, charged with providing the United States Congress non-partisan advice on issues that may come before Congress.

EveryCRSReport.com republishes CRS reports that are available to all Congressional staff. The reports are not classified, and Members of Congress routinely make individual reports available to the public.

Prior to our republication, we redacted names, phone numbers and email addresses of analysts who produced the reports. We also added this page to the report. We have not intentionally made any other changes to any report published on EveryCRSReport.com.

CRS reports, as a work of the United States government, are not subject to copyright protection in the United States. Any CRS report may be reproduced and distributed in its entirety without permission from CRS. However, as a CRS report may include copyrighted images or material from a third party, you may need to obtain permission of the copyright holder if you wish to copy or otherwise use copyrighted material.

Information in a CRS report should not be relied upon for purposes other than public understanding of information that has been provided by CRS to members of Congress in connection with CRS' institutional role.

EveryCRSReport.com is not a government website and is not affiliated with CRS. We do not claim copyright on any CRS report we have republished.