



State Assessments Required by the No Child Left Behind Act: An Analysis of Requirements, Funding, and Cost

name redacted

Analyst in Education Policy

March 3, 2011

Congressional Research Service

7-....

www.crs.gov

R41670

Summary

The Title I-A program of the Elementary and Secondary Education Act (ESEA), as amended by the No Child Left Behind Act (NCLB; P.L. 107-110), is the largest source of federal funding for elementary and secondary education. States that receive Title I-A funding must comply with certain requirements related to measuring student achievement by using state-level standards and assessments. The standards and assessment requirements adopted in the NCLB have resulted in vast growth in and demand for new types of assessment products. Estimates of annual expenditures on NCLB-required assessments range from \$500 million to \$900 million. The NCLB authorized a new grant assessment program to help states pay the additional costs of meeting the expanded assessment requirements. The purpose of this report is to provide an understanding of the factors that influence the overall cost of student assessment and a description of state expenditures on assessment systems.

This report describes factors that influence the cost of student assessment systems, including assessment development, scoring, and the effect of certain policy choices and implementation practices. It also reports and examines the results of existing research on state assessment expenditures, which highlight the variability among states in assessment expenditures. Next, the report presents the results of an analysis of how state assessment expenditures differ across states of varying populations. Results suggest that state assessment expenditures are dependent, in part, on a state's population; however, there are other factors that contribute to the overall cost of state assessment systems (e.g., design choices regarding assessment development and scoring).

This report examines assessment issues that Congress may address during the reauthorization of the ESEA. As Congress considers reauthorization, it is likely that the standards and assessment provisions will be reviewed and debated. State assessments are currently the primary policy tool used to measure whether schools, LEAs, and states are meeting certain academic targets. As such, the effectiveness of federal policy is dependent upon a valid state assessment system. Designing an assessment system requires states to consider certain tradeoffs in order to meet the federal assessment requirements while meeting state requirements and containing costs. Decisions made in the design of a state assessment system have direct implications for the validity of the federal accountability system and direct consequences for schools. This report examines assessment practices that may be cost-effective and could be promoted by federal policy, including the following:

- The use of various *scoring practices*—such as distributed scoring, teacher scoring, and artificial intelligence—that may have the potential to lower the cost of state assessment systems while maintaining quality.
- The use of *state consortia* that promote information sharing and other collaborative practices across states that may reduce assessment costs by realizing economies of scale.
- The use of *technology*, which may increase the efficiency of assessment administration. As technology becomes less expensive, online delivery of state assessments may lower administration and scoring costs.

Contents

Introduction	1
Assessment Requirements of the NCLB	2
General Assessment Requirements	2
Alternate Assessment Requirements for Students with Disabilities	4
Current Status of State Assessment Systems	4
The Assessment Industry.....	5
Estimated Costs for State Assessments	6
New Challenges for the Assessment Industry and States	7
Factors that Influence the Cost of State Assessment Systems	8
Assessment Development.....	10
Scoring	11
Policy Choices and Implementation Practices.....	12
State Expenditures on Assessment Systems Under the NCLB.....	13
U.S. Department of Education Study	13
U.S. Government Accountability Office Study	19
State Variability in Expenditures.....	21
Data Source and Questions for Analysis	22
Results	23
Promoting Efficiency in State Assessment Systems	25
Scoring Practices.....	26
State Consortia.....	28
Technology	29

Figures

Figure 1. State-Reported Aggregate Costs of Developing or Modifying Reading and Mathematics Assessments in Grades 3-8 to Comply with NCLB Requirements, in 27 States that Reported Such Data.....	15
Figure 2. State-Reported Annual Per-Pupil Costs of Administering and Maintaining Reading and Mathematics Assessments in Grades 3-8 and High School, 2006-2007, in 39 States that Reported Such Data.....	17
Figure 3. Average State-Reported Annual Total Cost and Per-Pupil Cost of Administering and Maintaining Reading and Mathematics Assessments in Grades 3-8 and High School, by State Enrollment Size, 2006-2007	18
Figure 4. State Expenditures for Assessment Vendors, 2007-2008.....	20
Figure 5. ESEA Assessment Activities That Received the Largest Share of States' Total ESEA Assessment Costs, 2007-2008	21

Tables

Table 1. Assessment Cost Structure	9
--	---

Table 2. Percentages of States' Funding for NCLB Assessments from Federal and State Government Funds, by State Size, 2007-2008..... 23

Table 3. State FTE Expenditures and Assessment Vendor Expenditures for NCLB Assessments, by State Size, 2007-2008 24

Table 4. Ratio of Assessment Vendor Expenditures to State FTE Expenditures, by State Size, 2007-2008 25

Table A-1. Federal Assessment Grants to States Under Section 6111: FY2002-FY2011 (estimate)..... 32

Appendixes

Appendix. Federal Funding for State Assessments..... 31

Contacts

Author Contact Information 34

Acknowledgments 34

Introduction

The Title I-A program of the Elementary and Secondary Education Act (ESEA), as amended by the No Child Left Behind Act (NCLB; P.L. 107-110), is the largest source of federal funding for elementary and secondary education. States that receive Title I-A funding must comply with certain requirements related to measuring and evaluating student achievement. Specifically, states are required to develop academic standards and assessments and use the results of these assessments within an accountability system that determines whether schools and local educational agencies (LEAs) are making adequate gains in student achievement.

Provisions regarding state standards and assessment systems were in place prior to the NCLB; they were enacted through the Improving America's Schools Act of 1994 (IASA; P.L. 103-382), which amended the ESEA. The IASA required states 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. Under the IASA, states were required to develop or adopt academic content standards, as well as content and performance standards and assessments tied to the content standards. The full system of standards and assessment was required to be in place by the 2000-2001 school year.¹

The standards and assessment provisions within the NCLB expanded upon provisions initially adopted in the IASA by increasing the number of grades and content areas that states are required to assess. Under the NCLB, states are required to assess students in grades 3 through 8 and one grade in high school in the content areas of reading and mathematics; states were required to administer these assessments by school year 2005-2006. In addition, states are required to assess students once in each of three grade bands (3-5, 6-8, and 10-12) in science; states were required to administer these assessments by school year 2007-2008.²

To help states pay the costs of meeting the expanded Title I-A requirements regarding standards and assessments, the NCLB authorized a new annual Grants for State Assessments program (Section 6111). The Grants for State Assessments program is a formula grant program that is intended to pay the costs of the development of the required assessments and, if assessments are fully developed, to help pay the costs of administering the assessments. In addition, the NCLB authorized a new annual Grants for Enhanced Assessment Instruments (Section 6112). The Grants for Enhanced Assessment Instruments program is a competitive grant program that is intended to enable states to improve the quality, validity, and reliability of their assessments beyond the requirements of Title I-A.

It is clear that state assessment expenditures have grown dramatically since the adoption of the NCLB. The extent to which these newly authorized grant programs have assisted states in meeting the assessment requirements of the NCLB is unclear. That is, it is unclear how much the federal grant programs are contributing to the overall cost of assessment activities under Title I-A of the NCLB. The primary purpose of this report is to explore issues related to the cost of

¹ By the end of the 2000-2001 school year, 16 states were granted full approval by ED for their assessment system under Title I of IASA. By the end of the 2003-2004 school year, 42 states were granted full approval. Decision letters on a state's final assessment system under Title I of IASA can be found at <http://www2.ed.gov/admins/lead/account/finalassess/index.html>.

² Because the NCLB introduced the current assessment requirements, "NCLB" is used throughout this report in lieu of "ESEA." In some places, "ESEA" is used if citing secondary sources or when discussing reauthorization issues.

assessments and to provide an understanding of the factors that influence the overall cost of assessment systems. This information is supplied to Congress as it prepares for the reauthorization of the ESEA and for deliberations on the ESEA standards and assessment provisions.

As Congress considers the reauthorization of the ESEA, it is likely that the standards and assessment provisions within Title I-A will be reviewed and debated. State assessments are currently the primary policy tool used to measure whether schools, LEAs, and states are meeting certain academic targets. As such, the effectiveness of federal policy is dependent upon a valid state assessment system. Designing an assessment system requires states to consider certain tradeoffs in order to meet the requirements of the NCLB while meeting state requirements and containing costs. Decisions made in the design of a state assessment system have direct implications for the validity of the NCLB accountability system and direct consequences for schools.³

This report aims to shed light on the current status of state assessment systems under the NCLB and it provides analysis that allows policymakers to explore some of the potential implications of changing the requirements of state assessment systems. The first section explicitly details the current assessment requirements of the NCLB and provides an overview of the current status of state assessment systems. The second section describes the assessment industry and how it responded to meet the expanded assessment requirements adopted in the NCLB. The third section provides a discussion of factors that influence the overall cost of student assessment within state assessment systems, which include assessment development, scoring and administration, policy choices, and implementation practices. The final sections report data on the cost of current assessment systems under the NCLB and explore some of the policies that may promote efficiency.

Assessment Requirements of the NCLB

Section 1111(b) outlines the state assessment requirements of the NCLB. Additional requirements for the assessment of students with disabilities are outlined in regulations. The general assessment requirements for all students and alternate assessment requirements for students with disabilities are discussed below. Following a discussion of the requirements is a summary of the current status of state assessment systems.

General Assessment Requirements

In general, each state is required to administer a set of high-quality, yearly academic assessments in mathematics, reading, and science. Results of these assessments are used to determine whether the state and each of its local educational agencies (LEAs) are assisting all students in meeting the achievement standards established by the state. In practice, the results of state assessments are used to determine whether schools and LEAs have made adequate yearly progress (AYP).⁴ In

³ For more information on the NCLB accountability system, see CRS Report RL33960, *The Elementary and Secondary Education Act, as Amended by the No Child Left Behind Act: A Primer*, by (name redacted).

⁴ For more information on AYP, see CRS Report RL32495, *Adequate Yearly Progress (AYP): Implementation of the No Child Left Behind Act*, by (name redacted).

addition, ESEA statutory provisions, as amended by the NCLB, specify that state assessments shall

- be used to measure the achievement of all students;
- be aligned with the state’s academic content and performance standards, and provide coherent information about student attainment of such standards;
- be used for purposes for which such assessments are valid and reliable, and be consistent with relevant, nationally recognized professional technical standards;
- be used only if the state educational agency (SEA) provides to the Secretary of Education (hereafter referred to as the Secretary) evidence from the test publisher or other relevant sources that the assessments used are of adequate technical quality for each purpose required by the ESEA;
- measure the proficiency in mathematics, reading or language arts, and science for grades 3 through 8 and one grade in high school;⁵
- involve multiple up-to-date measures of student academic achievement, including measures that assess higher-order thinking skills and understanding;
- provide for “reasonable” adaptations and accommodations for students with disabilities;
- provide for the inclusion of limited English proficient students, who shall be assessed, “to the extent practicable,” with assessments in the language that is most likely to yield accurate data on achievement;
- include students who have attended schools in an LEA for a full academic year;
- produce individual student reports that allow parents, teachers, and principals to understand the academic needs of the students, and include information regarding achievement aligned with state academic content standards in an understandable format as soon as is practicably possible;
- enable results to be disaggregated within each state, LEA, and school by gender, race/ethnicity, English language learner (ELL) status, migrant status, students with disabilities as compared to nondisabled students, and economically disadvantaged students as compared to students who are not economically disadvantaged;
- be consistent with accepted professional testing standards; objectively measure academic achievement, knowledge, and skills; and not evaluate or assess personal or family beliefs and attitudes or publicly disclose personally identifiable information; and
- enable itemized score analyses to be produced and reported to LEAs and schools so that parents, teachers, principals, and administrators can interpret and address

⁵ Mathematics and reading or language arts assessments were required to be administered beginning in school year 2005-2006. Science assessments were not required until school year 2007-2008. For more information on the required timing of the assessments, see ESEA, Section 1111(b)(3)(C)(v)-(vii).

the specific academic needs of students as indicated by the students' achievement on assessment items.⁶

Alternate Assessment Requirements for Students with Disabilities

In addition to the requirements above, states are required to develop at least one alternate assessment for students with disabilities.⁷ The requirement that states develop at least one alternate assessment was intended to ensure that all students with disabilities could participate in state assessment and accountability systems, which is required by both the NCLB and the Individuals with Disabilities Education Act (IDEA; P.L. 108-446). By the end of the 2005-2006 school year, all states had included at least one alternate assessment within their state assessment system.

There are currently five options for assessing students with disabilities for state accountability purposes: (1) general state assessment, (2) general state assessment with accommodations, (3) alternate assessment based on *grade-level* standards, (4) alternate assessment based on *alternate* achievement standards (AA-AAS), and (5) alternate assessment based on *modified* achievement standards (AA-MAS). The first three assessment options result in scores that may be counted in AYP calculations in the typical manner, as determined by a state's accountability system. Scores from the fourth and fifth assessment options (AA-AAS and AA-MAS) have restrictions on the way they may be counted in AYP calculations. These restrictions are outlined in regulations issued by ED and have numerous implications for state accountability systems.⁸

Current Status of State Assessment Systems

States have had varying degrees of success in developing state assessment systems that meet the expanded NCLB assessment requirements above. Although these requirements have been in place since the NCLB was enacted in 2002, 11 states have never received approval of their state assessment systems from the U.S. Department of Education (ED).⁹ As of January 2009, 39 states have received full approval for their reading and mathematics assessments, and 12 states are listed as "approval pending."¹⁰ In 11 of the 12 states, alternate assessment for students with

⁶ Requirements for state assessments under ESEA Title I are summarized here. For all of the requirements, see ESEA, Section 1111(b)(3)(C)(i)-(xv). In addition to requirements for state assessments under ESEA Title I, the Individuals with Disabilities Education Act (IDEA) requires that states develop at least one alternate assessment for students with disabilities. ED has developed regulations regarding the use of alternate assessments in state assessment and accountability systems. For more information on these regulations, see CRS Report R40701, *Alternate Assessments for Students with Disabilities*, by (name redacted).

⁷ The requirement that states develop alternate assessments for students with disabilities is not explicitly stated in statutory language; however, regulations on the development of these assessments have made the requirement explicit. For a comprehensive report on alternate assessments, see CRS Report R40701, *Alternate Assessments for Students with Disabilities*, by (name redacted).

⁸ See footnote 7.

⁹ In six of the 11 states, ED has entered into compliance agreements with the states, which allows the states an additional one to two years to come into compliance. ED has placed the remaining five states that never received approval into mandatory oversight status to determine if they have come into compliance. In addition, ED has placed conditions on FY2010 assessment grants in 29 states in which assessment systems have not yet received full approval.

¹⁰ The term "states" includes the 50 states and the District of Columbia. States with a status of "approval pending" have various agreements with ED for continuing their assessment systems (e.g., a memorandum of agreement or a compliance agreement).

disabilities was a factor preventing approval of the assessment system.¹¹ For science assessments, 10 states have received full approval.¹²

The Assessment Industry

With the adoption of the NCLB, there was a dramatic change in the demand for new assessments, which required the assessment industry to increase its production of assessment products. States began to require the assessment industry to develop new types of assessments that were reliable and valid for the purposes of their state accountability system. This section explores changes in the number and type of assessments that states required and how the industry and states have responded to these changes.

The assessment industry is a collection of private companies that develop, administer, and score assessments for a variety of purposes. Some companies in the assessment industry, also called “assessment vendors,” specialize in professional certifications, such as teacher certification or career and technical certifications. Others specialize in entry examinations for undergraduate or graduate programs. A subset of assessment vendors specialize in educational assessment for students in elementary and secondary schools.

The assessment requirements of the last two reauthorizations of the ESEA (i.e., the IASA and the NCLB) have greatly increased the demand for assessment development and administration services in elementary and secondary schools. Prior to the IASA, decisions regarding student assessment were largely made at the school or LEA level. Schools and LEAs typically purchased readily available, nationally normed tests of achievement, such as the Iowa Test of Basic Skills (ITBS) or the Stanford Achievement Test.¹³ Assessment vendors referred to these sales as “catalogue sales,” and the vendors could control development and administration costs fairly easily by using the same test for a number of years and across a number of LEAs.

The requirements of the IASA and the NCLB changed a fundamental element of assessment in elementary and secondary education—they required states to develop state academic content and performance standards and *align state assessments to the state content and performance standards*. This requirement in legislation caused a shift from the administration of norm-referenced assessments to the administration of criterion-referenced assessments that were aligned with state content and performance standards.¹⁴ Moreover, this requirement no longer

¹¹ In four states (Florida, New Hampshire, New Jersey, and Vermont), alternate assessment for students with disabilities was the only factor preventing approval.

¹² For more information on the status of state assessments, see U.S. Department of Education (ED), “State Standards and Assessment Update,” January 2009, <http://www2.ed.gov/admins/lead/account/statesystems.html>; U.S. Department of Education (ED), *FY2012 Budget Request*, Education Improvement Programs, pp. D-132, <http://www2.ed.gov/about/overview/budget/budget12/justifications/d-eip.pdf>.

¹³ For information on ITBS, see <http://www.riversidepublishing.com/products/itbs/index.html>. For information on the Stanford Achievement Test, see <http://www.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=SAT10C>.

¹⁴ For more information on the differences between norm-referenced and criterion-referenced assessments, see CRS Report R40514, *Assessment in Elementary and Secondary Education: A Primer*, by (name redacted). Note that states also had the option to “augment” existing norm-referenced assessments in order to align them to the state content and performance standards.

allowed states, LEAs, or schools to purchase assessments through “catalogue sales” because these assessments would not necessarily be aligned to the state content and performance standards.

The assessment requirements of the NCLB created a new market and increased demand for elementary and secondary assessments.¹⁵ Specifically, because of the requirement that assessments be aligned with content and performance standards, the assessment industry was tasked by states to create new, customized criterion-referenced assessments that measured each state’s unique content and performance standards. The relatively sudden demand for 52 new assessments¹⁶ created competition among assessment vendors to vie for state assessment contracts; in some cases, these state assessment contracts are worth hundreds of millions of dollars.

Although there is competition for state assessment contracts, there are a limited number of assessment vendors who develop the types of assessments required by the NCLB. It has been reported that five assessment vendors account for over 90% of the statewide testing revenue.¹⁷ There are several smaller, full-service assessment vendors and other niche companies that compete for state contracts as well.

Estimated Costs for State Assessments

In a conference report accompanying the passage of the NCLB, Congress required the Government Accountability Office (GAO) to conduct a study of the aggregate costs to develop and administer new state assessments between FY2002 and FY2008. GAO estimated that the aggregate costs of new state assessments would be between \$1.9 billion and \$5.3 billion, depending on the design of the assessments; this estimate represented the total cost to all states from FY2002 to FY2008. The low estimate of \$1.9 billion assumed that states would choose to develop only multiple-choice items that would be scored by machine. A mid-level estimate of \$3.9 billion assumed that states would maintain their current combination of multiple-choice and limited open-ended items as reported to GAO. The high estimate of \$5.3 billion assumed that states would use a combination of multiple-choice and open-ended items that require students to write their response, such as an essay, which would require human scoring.¹⁸

Since the passage of the NCLB, several nationwide estimates of expenditures on state assessments have been put forth. For school year 2005-2006, one estimate places nationwide assessment expenditures for NCLB-required state assessments at between \$500 million and \$750 million.¹⁹ A more recent estimate places nationwide expenditures at approximately \$900 million

¹⁵ See, for example, Karen Francisco, “No Dollar Left Behind: Federal Law Spawns Opportunity for Private Sector to Cash In,” *The Journal Gazette*, January 30, 2005; Kristin Proulx, “A Firm That Makes the Grade; Measured Progress Capitalizes on Growing Demand for Testing,” *Concord Monitor*, February 14, 2004; Tamar Lewin, “Corporate Culture and Big Pay Come to Nonprofit Testing Service,” *New York Times*, November 23, 2002, Late Edition.

¹⁶ “52 new assessments” presumes that each of the 50 states, the District of Columbia, and Puerto Rico would independently develop new assessments.

¹⁷ The five assessment vendors are Pearson Educational Measurement (Pearson); CTB/McGraw-Hill; Harcourt Assessment, Inc.; Riverside Publishing; and Educational Testing Service (ETS). For more information, see Thomas Toch, *Margins of Error: The Education Testing Industry in the No Child Left Behind Era*, Education Sector, Washington, DC, 2006, http://www.educationsector.org/sites/default/files/publications/Margins_of_Error.pdf.

¹⁸ U.S. General Accounting Office (GAO), *Title I: Characteristics of Tests Will Influence Expenses; Information Sharing May Help States Realize Efficiencies*, GAO-03-389, May 2003, <http://www.gao.gov/new.items/d03389.pdf>.

¹⁹ For more information, see Thomas Toch, *Margins of Error: The Education Testing Industry in the No Child Left* (continued...)

per year.²⁰ Although the assumption that states would maintain their pre-NCLB combination of multiple-choice and limited open-ended items has not been investigated, recent estimates of expenditures suggest that the original GAO estimates were low.

New Challenges for the Assessment Industry and States

The assessment industry and states faced new challenges in implementing the assessment requirements of the NCLB. One observer noted that “the massive expansion of required testing fell onto the relatively small testing industry like a log landing on a twig.”²¹ Although there was at least one report asserting that the assessment industry was prepared to meet the new demand,²² the media often reported that assessment vendors faced difficulties in increasing staff capacity to handle the workload. Some have reported that the assessment industry is “stretched thin”²³ and that vendors are “competing against each other for the same people.”²⁴ One particular area of concern is a reported shortage of measurement experts, or psychometricians, to assess the reliability and validity of assessments. It has been estimated that between 1995 and 2003, under a dozen doctoral degrees were awarded to psychometricians each year; an additional 35 doctoral degrees were granted annually in the fields of statistics, testing, and educational measurement.²⁵ The educational assessment industry must compete with other industries for a limited number of psychometricians, leading to the phenomenon of “psychometrician musical chairs.”²⁶

There is some evidence that the assessment industry was not prepared to meet the new demand created by the NCLB. Numerous errors in printing, administering, and scoring assessments have been reported.²⁷ In some cases, assessment results were delayed by several months.²⁸ A survey of 23 states found that 35% of states experienced significant errors with scoring, and 20% of states reported delays in receiving assessment scores.²⁹ States have penalized some assessment vendors for the errors and delays through fines or contract termination.³⁰

(...continued)

Behind Era, Education Sector, Washington, DC, 2006, http://www.educationsector.org/sites/default/files/publications/Margins_of_Error.pdf.

²⁰ E-mail from assessment vendor, personal contact, July 7, 2010.

²¹ Jay Matthews, “Assessment Industry Faces a Test of Its Own: Pushed by “No Child” and Plagued by Software Glitches, Companies Face Strains That Could Reshape Exams,” *The Washington Post*, May 28, 2007, Final Edition.

²² Stuart R. Kahl, *NCLB Testing: Is the Assessment Industry Up for the Challenge?*, Measured Progress, Issue Paper Number 1, Dover, NH, October 2003.

²³ Holly Prestidge, “Pearson Not Alone in SOL Mistakes; Ex-Provider Harcourt Has Scored Incorrectly in VA and Elsewhere,” *Richmond Times-Dispatch*, October 27, 2005, City Edition.

²⁴ Mark Fisher and Scott Elliott, “Flunking the Test,” *Dayton Daily News*, May 23, 2004, City Edition.

²⁵ See footnote 19.

²⁶ See footnote 21.

²⁷ See, for example, Mark Fisher, “Flunking the Test,” *Dayton Daily News*, May 23, 2004, City Edition; Robert A. Frahm, “State Fires Testing Contractor; Signs 5-year Pact with Different Firm,” *Hartford Courant*, September 13, 2005, Final Edition; and Megan Reichgott, “No Child Taxing Testing Process: Delays, Errors Rise as Exam Companies Struggle to Keep Up,” *Associated Press*, March 26, 2007.

²⁸ Pete Sherman, “Schools’ Assessment Test Scores Will Be Late,” *State Journal Register*, January 22, 2007.

²⁹ See footnote 19.

³⁰ See, for example, Robert A. Frahm, “State Fires Testing Contractor; Signs 5-year Pact with Different Firm,” *Hartford Courant*, September 13, 2005, Final Edition; Maria Garriga, “Maximum Fine Sought for Test Scoring Flub,” *New Haven Register*, February 17, 2006; and Holly Prestidge, “Pearson Not Alone in SOL Mistakes; Ex-Provider (continued...)”

States have also faced challenges in implementing the new assessment requirements of the NCLB. State assessment offices were charged with administering competitions for new assessment contracts, which often required an “enormous vetting process,” including reviewing proposals and presentations from vendors.³¹ The proposals often must be vetted by a panel of teachers, school district testing directors, data management specialists, and other interested stakeholders.³² Some states have hired new staff members to manage the workload; however, new state employees are “on a learning curve in terms of figuring out how to collect the data, clean it, and turn around and report it back.”³³ A survey of state testing directors found that over half of the states have problems recruiting and retaining key staff needed to respond to administering the NCLB. Reportedly, there is also a high turnover rate among testing directors, many of whom have left state jobs to pursue opportunities in the private sector.³⁴

Even when states have the necessary capacity to oversee assessment contracts, some states have reported that it is difficult to remedy situations in which vendors have caused extensive delays or made repeated errors. Other than the threat of financial penalties, the states report that they have little leverage to cancel or modify contracts because if the contracts were re-competed, “the same companies would bid.”³⁵ To this end, one testing expert noted that the assessment industry is “not a monopoly, but it is an oligopoly, with very little regulation.”³⁶

Factors that Influence the Cost of State Assessment Systems

There is insufficient research on factors that influence the cost of assessments required by the NCLB to construct a complete model of overall cost. Much of the research on cost predates the NCLB, and it may not provide accurate depictions of assessment practices today.³⁷ In lieu of existing information regarding cost, this section of the report presents a framework to guide discussion about the factors that influence the cost of student assessment.³⁸ It focuses on

(...continued)

Harcourt Has Scored Incorrectly in VA and Elsewhere,” *Richmond Times-Dispatch*, October 27, 2005, City Edition.

³¹ Cara Fitzpatrick, “FCAT Errors Don’t Faze States; Florida Plans to Give More Work to a Firm with a List of Blunders,” *Palm Beach Post*, June 27, 2010.

³² Duke Helfand, “ETS to Continue State’s Testing,” *Los Angeles Times*, November 11, 2005, Home Edition.

³³ Susan Saulny, “Delayed Test Results Trouble New York,” *New York Times*, August 12, 2005, Late Edition (East Coast).

³⁴ See footnote 19.

³⁵ See footnote 31.

³⁶ Megan Reichgott, “No Child Taxing Testing Process: Delays, Errors Rise as Exam Companies Struggle to Keep Up,” *Associated Press*, March 26, 2007.

³⁷ See, for example, Lawrence O. Picus, *A Conceptual Framework for Analyzing the Costs of Alternative Assessment*, National Center for Research on Evaluation, Standards, and Student Testing, CSE Technical Report 384, Los Angeles, CA, August 1994, <http://www.cse.ucla.edu/products/Reports/TECH384.pdf>; Lawrence O. Picus, *Estimating the Costs of Student Assessment in North Carolina and Kentucky: A State Level Analysis*, National Center for Research on Evaluation, Standards, and Student Testing, CSE Technical Report 408, Los Angeles, CA, February 1996, <http://www.cse.ucla.edu/products/Reports/TECH408.pdf>.

³⁸ The factors discussed in this section include assessment design issues and policy choices that affect the cost of assessment. This section does not discuss costs related to student population. The relationship between assessment costs and student population is discussed in a later section, “State Variability in Expenditures”.

assessment development, scoring, policy choices, and implementation practices. After these components of state assessment systems are introduced, the research that exists on costs associated with these components is discussed. This is followed by an original analysis which, within limitations imposed by available data, tests the assumption that student population is a major factor affecting cost.

Table 1. Assessment Cost Structure

Proportional costs of various assessment components: low, mid-level, and high estimates

Factor	Low	Mid-level	High
Development	18%	22%	25%
Production & manufacturing	12%	15%	18%
Warehousing & shipping	12%	13%	15%
Document preparation, scanning, editing, multiple choice scoring	12%	15%	18%
Open-ended scoring	12%	16%	20%
Reporting	2%	4%	5%
Information technology, quality assurance, and other	12%	15%	18%

Source: E-mail from assessment vendor, personal contact, July 15, 2010.

Table 1 provides a cost structure that outlines low, mid-level, and high estimates for various assessment components. For example, development accounts for approximately 18% to 25% of the total assessment cost, depending on a state's relative level of investment in this activity. This cost structure can be used to assist states in balancing their assessment budget; it is designed to help states understand the tradeoffs involved in creating an assessment system. At the mid-level, all of the components sum to 100%. If, however, a state chooses to spend a "high" proportion of the assessment budget on development (moving from 22% to 25%), it would need to off-set that expense by choosing to spend a "low" proportion of the budget on a component like production and manufacturing. Any combination that sums to 100% would represent a balanced assessment budget.

Across the low, mid-level, and high estimates, development is the most expensive assessment component, accounting for 18% to 25% of total assessment cost. The least expensive assessment component in the cost structure is reporting, accounting for 2% to 5% of the total assessment cost. All other components are relatively equal in terms of their proportional cost, ranging anywhere from 12% to 20% of the total assessment cost.

Some of the components in this cost structure are relatively fixed costs that would be similar regardless of the type of assessment used (e.g., reporting); however, other components are assessment design elements over which states can exert some control (e.g., development; document preparation, scanning, editing, multiple choice scoring; open-ended scoring). The sections below discuss two of the major design elements that can influence costs: assessment development and scoring. The final section discusses various policy choices and implementation practices that can also influence costs.

Assessment Development

Assessment development is an ongoing process. Test developers often convene to create a new set of assessment items based on academic content and performance standards used in the state. Once created, the items are field-tested to allow test developers and psychometricians to investigate the reliability and validity of the assessment. If certain items prove to be unreliable or invalid, they are removed. Once problematic items are removed, the remaining items are sometimes reviewed by committees to check for potential bias against certain student groups. At this point, the assessment may be administered to students for the purpose of accountability. Once administered for the purpose of accountability, it is common practice for the state to release some of the assessment items to the public. This practice is used to build credibility and create buy-in to the overall assessment and accountability system of the state. The items that are released to the public must be replaced annually.

The cost of assessment development is primarily dependent on the number of assessments to be developed and the types of assessment to be developed. The number of assessments developed by states to fulfill the requirements of the NCLB is, at minimum, 17. The 17 assessments include reading assessments for grades 3 through 8 and one grade in high school, mathematics assessments for grades 3 through 8 and one grade in high school, and science assessments for three grade levels.³⁹

The most common types of assessment used by states to fulfill the requirements of the NCLB are (1) multiple choice assessments, or (2) multiple choice assessments with constructed response.⁴⁰ Multiple choice items require students to choose the correct answer from a selection of responses; constructed response items require students to generate a written response.⁴¹ Since 2002, states have generally reported increasing the use of multiple choice items and decreasing the use of constructed response items.⁴²

The cost of developing a single multiple choice item has been estimated to be between \$300 to \$1,000. State assessments typically include 50 to 100 items per subject (i.e., reading, mathematics, science) per grade.⁴³ Using these parameters, a conservative estimate for assessment development, therefore, would be around \$255,000.⁴⁴ It is possible, however, that assessment development could be as high as \$1.7 million.⁴⁵ These estimates do not include assessment

³⁹ In addition to assessments in reading, mathematics, and science, states must also develop at least one alternate assessment for students with disabilities. Because states have discretion in the number and type of alternate assessments developed, it is difficult to include them in an overall discussion of required assessments. It is likely, however, that 17 assessments is a conservative depiction of the number of assessments that states develop.

⁴⁰ U.S. General Accounting Office (GAO), *Characteristics of Tests Will Influence Expenses; Information Sharing May Help States Realize Efficiencies*, GAO-03-389, May 2003.

⁴¹ Constructed response items vary with regard to the amount of written information required. Some constructed response items require the student to supply a specific term or demonstrate how to solve a math problem. Other extended constructed response items may require the student to write an essay or demonstrate a mathematical proof.

⁴² U.S. Government Accountability Office (GAO), *Enhancements in the Department of Education's Review Process Could Improve State Academic Assessments*, GAO-09-911, September 2009.

⁴³ Information on the cost of multiple choice items and the number of items per state assessment can be found in Thomas Toch, *Margins of Error: The Education Testing Industry in the No Child Left Behind Era*, Education Sector, Washington, DC, 2006, http://www.educationsector.org/sites/default/files/publications/Margins_of_Error.pdf.

⁴⁴ This estimate assumes 17 assessments, 50 items per assessment, and \$300 per item.

⁴⁵ This estimate assumes 17 assessments, 100 items per assessment, and \$1,000 per item.

development costs associated with any constructed response items. Furthermore, they do not include the ongoing costs of developing new items each year to replace those items that were released to the public. Therefore, estimates presented herein may underestimate the total cost of development.

Scoring

Scoring assessments includes both automated scoring and human scoring. Automated scoring by machine is used for multiple choice items and some constructed response items. Human scoring is used for constructed response items that require judgment regarding the correctness or quality of the response. Typically, human scoring is conducted by staff hired by the assessment vendor that is administering the assessment.

Automated scoring is relatively fast and efficient, and human scoring is relatively slow and laborious. Human scoring requires extensive training and oversight of scorers in order to ensure a basic standard of accuracy and reliability. For example, it is common for scorers to participate in one or more training sessions in order to establish a reliability with the scoring procedures and other scorers in the group. Next, assessments are distributed for scoring. Most times, a certain percentage of assessments are sent to more than one scorer for “double scoring,” sometimes called a “read behind.” The read-behind procedure allows an assessment vendor to determine whether the reliability between the scorers is high enough to ensure the accuracy and validity of the results.

Since most states use multiple choice assessments to meet the assessment requirements of the NCLB, the majority of the scoring is automated. The speed and ease of automated scoring may actually be a strong incentive for states to use multiple choice assessments. States face a certain time pressure to report results quickly and accurately. NCLB regulations require states to report the results of state assessments before the beginning of the next school year.⁴⁶ Generally, states would prefer to conduct assessments as late in the school year as possible to reflect the achievement level of students after a year of instruction. If assessments are administered in late spring and results are needed by mid-summer, states may only be afforded a window of several weeks to score assessments.

Some states have expressed interest in using more constructed response items in their state assessment systems under the NCLB. One reason states may choose to use multiple choice items over constructed response items is the associated scoring costs. In 2003, GAO reported data on the relative costs of scoring an assessment that primarily used multiple choice items versus an assessment that used multiple choice and constructed response.⁴⁷ In the states of North Carolina and Virginia, assessments primarily consist of multiple choice items; in these states, scoring costs less than \$1 per assessment. Massachusetts, on the other hand, uses an assessment with both multiple choice and constructed response items; scoring costs approximately \$7 per assessment.

⁴⁶ Regulations specify that an LEA must notify parents of public school choice options made available to students who attend a school in need of improvement. Parental notification “must be made sufficiently in advance of, but no later than 14 calendar days before, the start of the school year so that parents have adequate time to exercise their choice option before the school year begins.” See 34 C.F.R. §200.37.

⁴⁷ In the GAO report, constructed response items are referred to as “open-ended.” For the purposes of this report, these terms are interchangeable.

Another reason states choose multiple choice over constructed response items is the extended timeline required by human scoring. The current window to score assessments (i.e., several weeks) may not allow for large-scale assessment programs with human scoring. Given this short turnaround time, states may opt for speed, accuracy, and cost effectiveness over the potential benefits of constructed response items that require human scoring.⁴⁸

Policy Choices and Implementation Practices

Federal and state policies set the requirements for state assessments and influence the design of the overall assessment system. As discussed, the number of assessments required by federal policy can influence the cost of assessments. Additional state policies can increase the cost of the assessment system as well. In addition to the assessment requirements of the NCLB, states sometimes require assessments in additional grades or subject areas. Furthermore, some state policies require developing multiple versions of each assessment to reduce the likelihood of cheating or for use in cases of student absence. For example, GAO reported in 2003 that Massachusetts used 24 different versions for many of its assessments and spent approximately \$200,000 to develop each assessment. By contrast, Texas used one version of its assessment and spent approximately \$60,000 to develop each assessment.⁴⁹

State policies also vary with respect to releasing items to the public. GAO reported that 54% of states release items to the public; however, the number of items released each year is highly variable. Texas, Massachusetts, Maine, and Ohio released the entire assessment to the public to allow parents and other stakeholders to see every test item administered to students. New Jersey and Michigan reported releasing only a portion of their assessments. Some states, such as North Carolina, reported releasing items to teachers but not the public.⁵⁰ Due to the variability in the number of items released each year, states face different replacement rates for assessment items. States that release the entire assessment annually must redevelop the assessment annually. States that release only a portion of the assessment have fewer items to redevelop annually. Even states that do not release items to the public replace some items each year for the purpose of test security. State policies regarding the number of items released to the public have direct consequences on the development costs in a state assessment system.

Another state policy that influences the cost of an assessment system is participation in a state consortium. Currently, there is one consortium of states that works to develop, maintain, and administer assessments that fulfill the requirements of the NCLB. The New England Common Assessment Program (NECAP) is a collaborative endeavor that includes Maine, New Hampshire, Vermont, and Rhode Island. Maine was the most recent state to join the consortium, and state officials have reported saving \$1 million per year on state assessments as part of the NECAP consortium.⁵¹

⁴⁸ Interactions between innovative assessment, scoring, timeline, and cost are discussed in a later section, “Promoting Efficiency in State Assessment Systems”.

⁴⁹ See footnote 40.

⁵⁰ See footnote 40.

⁵¹ Kelley Bouchard, “Regional Tests Will Save Maine \$1 Million a Year,” *Portland Press Herald*, December 28, 2008.

State Expenditures on Assessment Systems Under the NCLB

The previous section explored the cost structure of state assessments and how some of the factors of a state assessment system influence cost. The current section turns to a discussion of how states report spending money on state assessment systems required by the NCLB.

There are relatively little data available that reflect state expenditures on their state assessment activities required by the NCLB. One reason for the lack of data is the difficulty in accounting for assessment expenses across multiple, and sometimes overlapping, activities. For example, states typically use assessment vendors to develop, administer, and score assessments. These vendors, however, may perform other duties that are not directly related to assessment (such as developing or maintaining standards) or perform other assessment duties that are not directly related to assessments required by the NCLB (such as developing end-of-course assessments or developing assessments for additional grades or content areas that are required by the state but not the NCLB). In cases where assessment vendors conduct additional work within one contract, it is difficult for states to separately account for expenses directly related to assessment activities required by the NCLB.

In addition to awarding contracts to assessment vendors, states also employ state- and local-level assessment directors or coordinators. In some cases, states have full-time staff, including directors, coordinators, and measurement specialists, who are hired primarily to handle assessment activities. In other cases, state- and local-level officials have overlapping duties and are not hired specifically to direct or coordinate assessments. For example, some states or LEAs may have a director of Title I programs, and one of the duties may be coordinating assessments and managing the contracts awarded to assessment vendors. The assessment duties may actually be a small part of their job, and states may not be able to separately account for the portion of salaries spent on assessment versus other duties.

Because of these accounting difficulties, reports of state expenditures on assessment activities required by the NCLB are highly variable. Two recent reports have investigated states' reports of costs associated with complying with the NCLB assessment requirements. Results of each investigation are discussed below along with a discussion of factors that may contribute to the variability reported across states.

U.S. Department of Education Study⁵²

As part of the National Assessment of Title I, Congress required an examination of the cost of developing assessments in grades 3 through 8.⁵³ ED interpreted this requirement as pertaining to general assessments in reading and mathematics. ED did not examine the cost of developing

⁵² James Taylor, Brian Stecher, and Jennifer O'Day, et al., *State and Local Implementation of the No Child Left Behind Act*, U.S. Department of Education (ED), Volume IX—Accountability Under NCLB: Final Report, Washington, DC, 2010, <http://www2.ed.gov/rschstat/eval/disadv/nclb-accountability/nclb-accountability-final.pdf>.

⁵³ ESEA, Section 1501(a)(2)(C)(i). Although the ESEA requires that academic assessments are administered at one grade in high school, the National Assessment of Title I only required ED to examine the cost of academic assessments in grades 3 through 8.

assessments in science, nor did it examine the cost of developing alternate assessments for students with disabilities.⁵⁴ As such, reports of state expenditures on assessment activities that were obtained by ED likely underestimate the true expenditures for the state assessment system.

ED surveyed state assessment directors and asked them to provide information on the cost of their state assessment program for school year 2006-2007. State directors were instructed to report cost estimates for assessment *development* and assessment *administration and maintenance*. For reporting development costs, officials were asked to aggregate costs over time because test development is considered a multiyear process. For reporting administration and maintenance costs, officials were asked to provide estimates of costs only for school year 2006-2007.

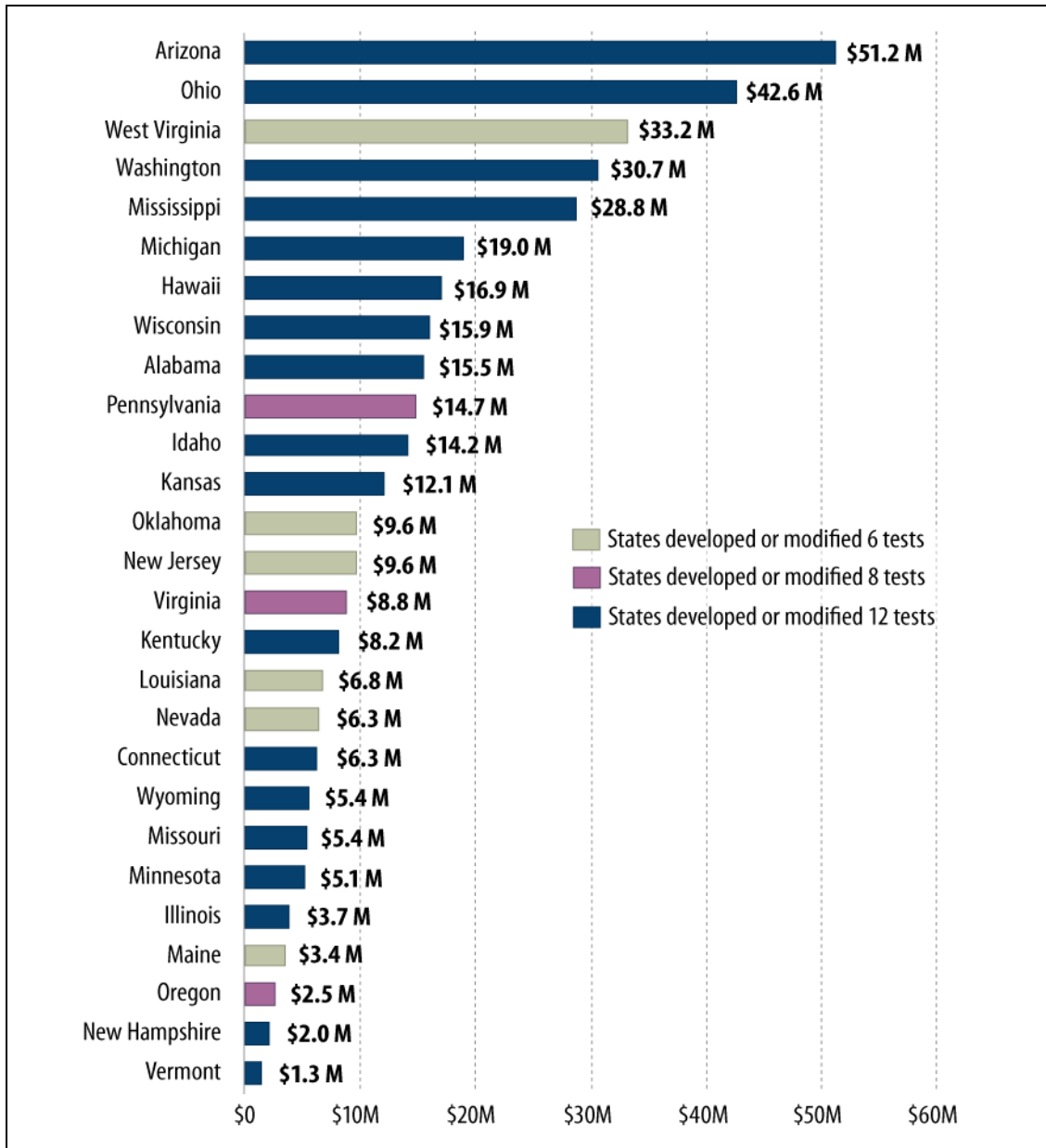
Twenty-seven states reported the cost of developing or modifying reading and mathematics assessments to comply with the NCLB. Of the states that reported data, three states reported developing or modifying six assessments; six states reported developing or modifying eight assessments; and 18 states reported developing or modifying 12 assessments. **Figure 1** shows the state-reported aggregate costs of developing or modifying reading and mathematics assessments in grades 3 through 8.

The average state-reported aggregate cost was approximately \$14.0 million; the median aggregate cost was \$9.6 million. State-reported aggregate costs, however, were highly variable. The level of state-reported costs for assessments ranged from a low of approximately \$1.3 million to a high of approximately \$51.2 million. Moreover, the variability does not appear to be related to the number of assessments that were developed or modified, nor does it appear to be related to state population. For example, in terms of the number of assessments that were developed or modified, the state with the lowest reported costs (Vermont) and the state with the highest reported costs (Arizona) both developed or modified 12 assessments, indicating that they changed every reading and mathematics assessment in grades 3 through 8.⁵⁵ In terms of state population, a relatively small state (West Virginia) reported spending over 10 times the amount of a relatively large state (Illinois). Furthermore, Illinois reported developing or modifying a greater number of assessments than West Virginia, making the relationship between population, number of assessments, and costs even more ambiguous.

⁵⁴ Because science assessments were not required to be in place until school year 2007-2008, ED surveyed state assessment directors during the 2006-2007 school year; therefore, the costs of developing science assessments are not reflected in ED's survey data.

⁵⁵ Vermont is viewed by some as an outlier in state assessment costs because it participates in the NECAP consortium of states to develop and maintain a single assessment system for multiple states. Nevertheless, there were five other states that developed or modified 12 assessments and reported costs below \$6.3 million; therefore, the number of assessments that were developed or modified does not appear to directly explain the overall costs incurred, as reported.

Figure I. State-Reported Aggregate Costs of Developing or Modifying Reading and Mathematics Assessments in Grades 3-8 to Comply with NCLB Requirements, in 27 States that Reported Such Data



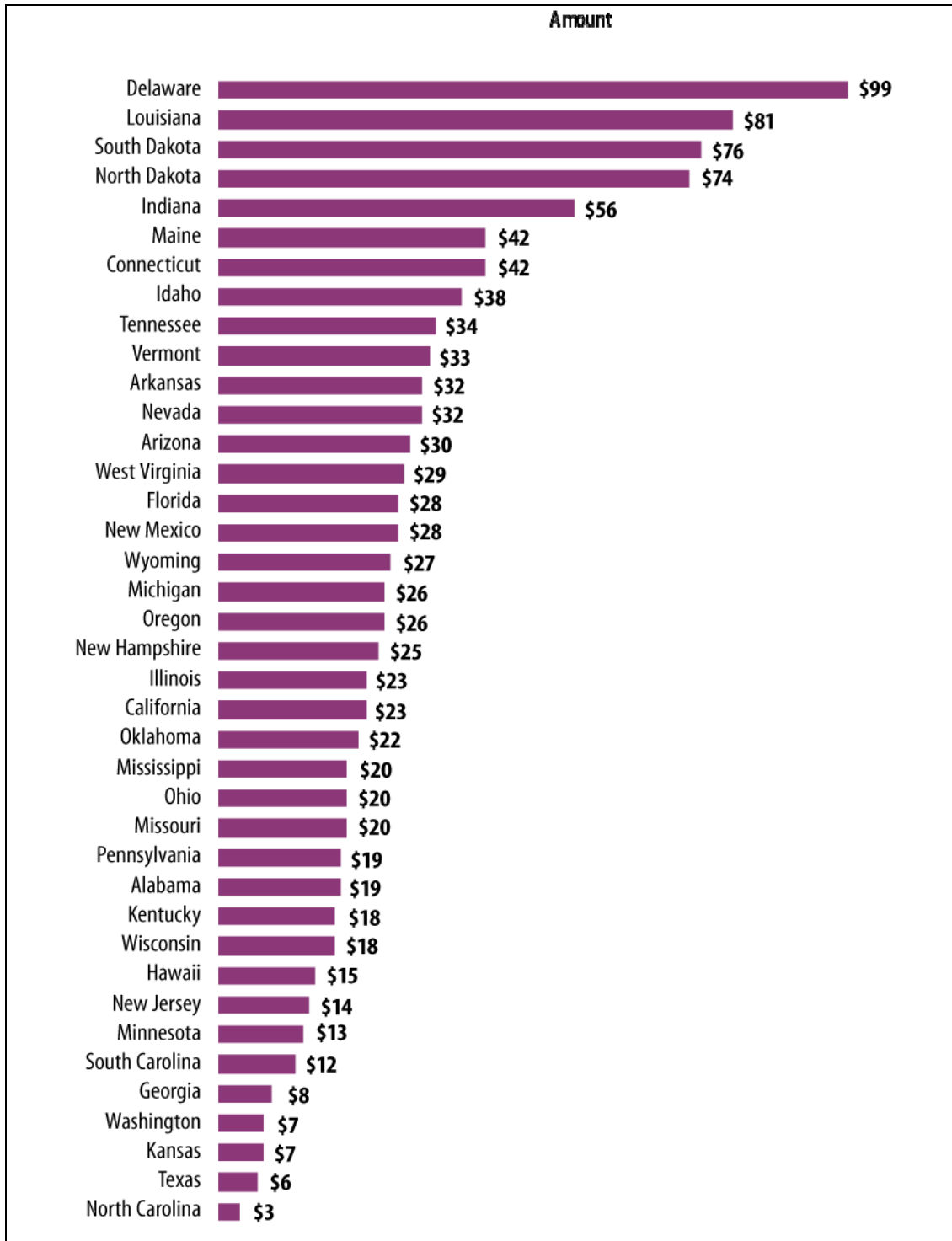
Source: Figure reproduced by CRS from data reported by the U.S. Department of Education. See James Taylor, Brian Stecher, and Jennifer O'Day, et al., *State and Local Implementation of the No Child Left Behind Act*, U.S. Department of Education (ED), Volume IX—Accountability Under NCLB: Final Report, Washington, DC, 2010, <http://www2.ed.gov/rschstat/eval/disadv/nclb-accountability/nclb-accountability-final.pdf>.

Thirty-nine states reported annual per-pupil costs of administering and maintaining reading and mathematics assessments in grades 3 through 8 and one grade in high school for school year 2006-2007.⁵⁶ **Figure 2** shows the state-reported annual, per-pupil costs.

The average state-reported per-pupil cost was approximately \$29; the median per-pupil cost was \$25. State-reported annual, per-pupil costs were highly variable. The annual per-pupil cost of administering and maintaining assessments ranged from a low of \$3 (North Carolina) to a high of \$99 (Delaware).

⁵⁶ It is unclear why the ED survey required states to include “aggregate” costs for grades 3 through 8 only while they required states to include “per-pupil” costs for grades 3 through 8 and one high school grade. It is possible that the survey was designed to be responsive to typical state accounting practices to encourage a higher survey response.

Figure 2. State-Reported Annual Per-Pupil Costs of Administering and Maintaining Reading and Mathematics Assessments in Grades 3-8 and High School, 2006-2007, in 39 States that Reported Such Data

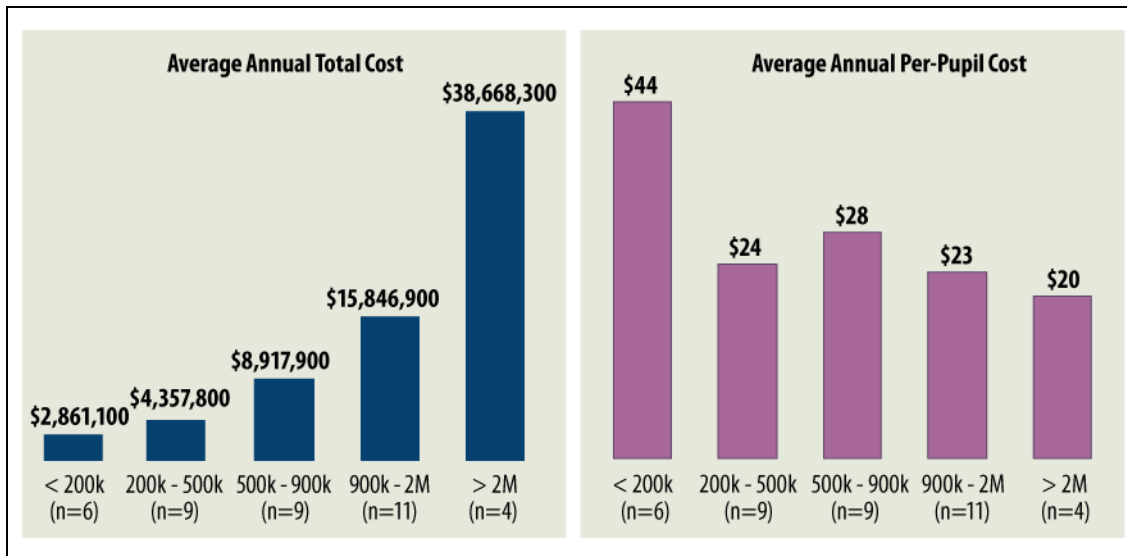


Source: Figure reproduced by CRS from data reported by the U.S. Department of Education. See James Taylor, Brian Stecher, and Jennifer O'Day, et al., *State and Local Implementation of the No Child Left Behind Act*, U.S.

Department of Education (ED), Volume IX—Accountability Under NCLB: Final Report, Washington, DC, 2010, <http://www2.ed.gov/rschstat/eval/disadv/nclb-accountability/nclb-accountability-final.pdf>.

Figure 3 shows the relationship between (1) average state-reported annual costs of administering and maintaining the reading and mathematics assessments and state enrollment, and (2) average state-reported per-pupil costs of administering and maintaining the reading and mathematics assessments and state enrollment. Larger states reported higher annual total costs for administering and maintaining assessments. ED reported that the costs for the four states with the highest student enrollment were higher than the six states with the lowest student enrollment (\$38.7 million compared to \$2.9 million). Smaller states, however, generally reported higher per-pupil costs for administering and maintaining assessments than larger states. ED reported that the average costs for the four states with the highest student enrollment were lower than those for the states with the lowest student enrollment (\$20 compared to \$44). While larger states reported higher costs for administering and maintaining assessments, it is likely that they benefitted from economies of scale that led to lower per-pupil costs.

Figure 3. Average State-Reported Annual Total Cost and Per-Pupil Cost of Administering and Maintaining Reading and Mathematics Assessments in Grades 3-8 and High School, by State Enrollment Size, 2006-2007



Source: Figure reproduced by CRS from data reported by the U.S. Department of Education. See James Taylor, Brian Stecher, and Jennifer O'Day, et al., *State and Local Implementation of the No Child Left Behind Act*, U.S. Department of Education (ED), Volume IX—Accountability Under NCLB: Final Report, Washington, DC, 2010, <http://www2.ed.gov/rschstat/eval/disadv/nclb-accountability/nclb-accountability-final.pdf>.

U.S. Government Accountability Office Study⁵⁷

From August 2008 through September 2009, GAO conducted a performance audit to address the following questions: (1) How have state expenditures on assessments required by the ESEA changed since enactment of the NCLB in 2002, and how have states spent funds? (2) What factors have states considered in making decisions about item type and the content of their assessments? (3) What challenges, if any, have states faced in ensuring the validity and reliability of their assessments? (4) To what extent has ED supported and overseen state efforts to comply with the NCLB assessment requirements? To address these questions, GAO used a mixed-methods research approach that included reviews of state documents, a web-based survey of the 50 states and the District of Columbia, interviews with ED officials and assessment experts, site visits to four states, and a review of federal laws and regulations. For the purpose of this report, the results reported herein focus on GAO's first two questions, which are directly related to state assessment expenditures.

Forty-nine of 51 states responded to the web-based survey.⁵⁸ Forty-eight of the 49 states reported that the state's annual assessment expenditures have increased since enactment of the NCLB. Over half of the 48 states reported that the primary reason for increased expenditures was the requirement to develop and administer additional assessments.⁵⁹ The second most commonly cited reason for increased assessment expenditures was increased assessment vendor costs.

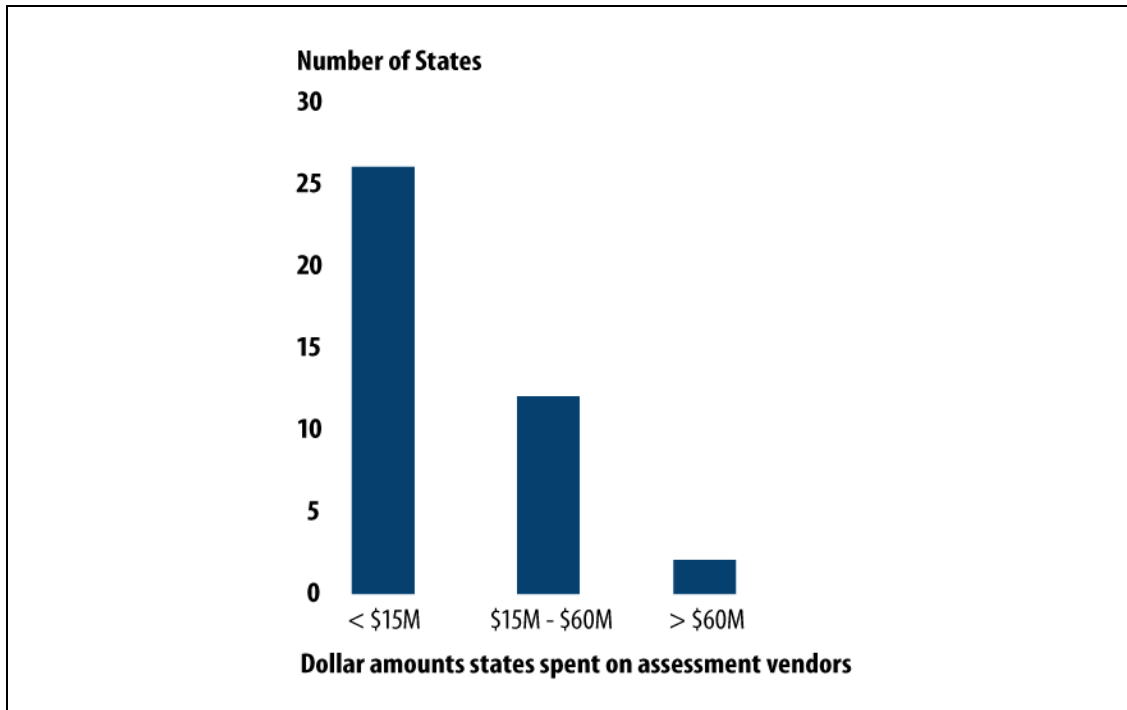
Figure 4 shows state assessment expenditures of 46 states for assessment vendors during school year 2007-2008. Of the 46 states that responded to the survey item, 44 states reported higher expenditures for assessment vendors than for state assessment staff. Of the 20 states that provided information on actual costs, GAO estimated that states spent over 10 times more on assessment vendors than on state assessment staff.

⁵⁷ U.S. Government Accountability Office (GAO), *No Child Left Behind Act: Enhancements in the Department of Education's Review Process Could Improve State Academic Assessments*, GAO-09-911, September 2009, <http://www.gao.gov/new.items/d09911.pdf>.

⁵⁸ New York and Rhode Island did not respond.

⁵⁹ Prior to the NCLB, states were not required to test students at as many grade levels.

Figure 4. State Expenditures for Assessment Vendors, 2007-2008



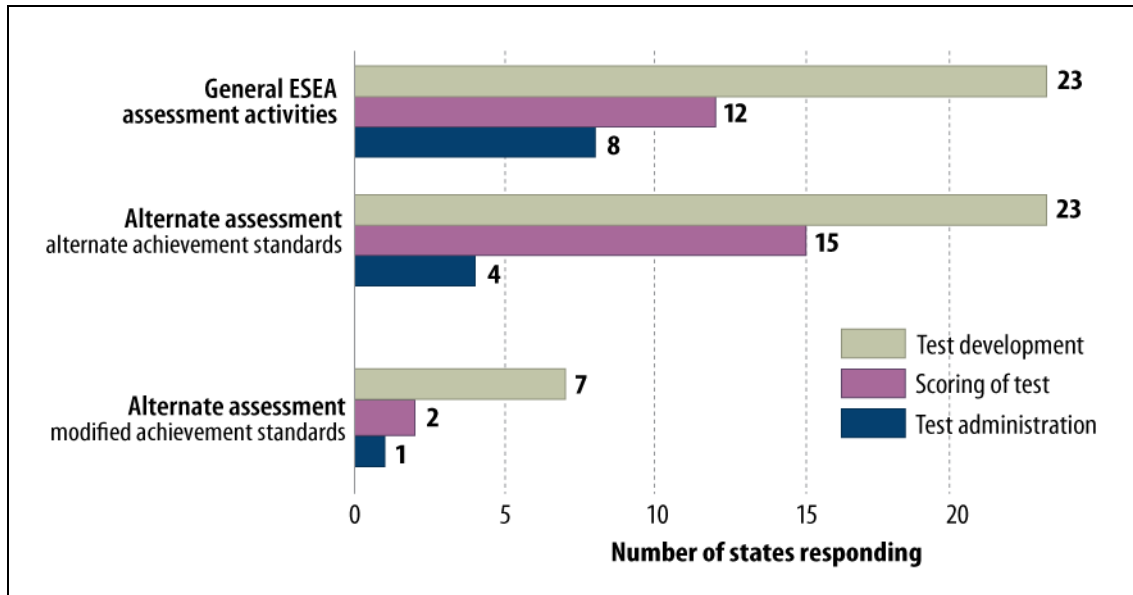
Source: Figure reproduced by CRS from data provided by the U.S. Government Accountability Office. See U.S. Government Accountability Office (GAO), *No Child Left Behind Act: Enhancements in the Department of Education’s Review Process Could Improve State Academic Assessments*, GAO-09-911, September 2009, <http://www.gao.gov/new.items/d09911.pdf>.

States were asked to rank assessment components in terms of each component’s relative cost.⁶⁰ **Figure 5** shows the NCLB assessment activities that received the highest proportion of state assessment expenditures for school year 2007-2008. For the general state assessment,⁶¹ 23 of 43 states ranked test development as the most expensive assessment activity; 12 states ranked scoring as the most expensive assessment activity; and eight states reported that administration was the most expensive assessment activity. GAO reported that one reason for high test development expenditures is that most states release a certain number of their assessment items to the public each year. As discussed, the released items need to be replaced with newly developed items each year. States vary greatly with respect to the proportion of items they release. Nearly all states release some items in order to build credibility with policymakers and the public, but a handful of states may release 20%, 50%, or even 100% of their items each year.

⁶⁰ States were asked to rank the following assessment components: test development, scoring, administration, reporting, data management, and “all other assessment activities.”

⁶¹ “General state assessment” includes reading and mathematics assessments in grades 3 through 8 and one grade in high school; it does not include alternate assessments used for some students with disabilities.

Figure 5. ESEA Assessment Activities That Received the Largest Share of States' Total ESEA Assessment Costs, 2007-2008



Source: Figure reproduced by CRS from data provided by the U.S. Government Accountability Office. See U.S. Government Accountability Office (GAO), *No Child Left Behind Act: Enhancements in the Department of Education's Review Process Could Improve State Academic Assessments*, GAO-09-911, September 2009, <http://www.gao.gov/new.items/d09911.pdf>.

The ranking pattern for the most expensive assessment activities was consistent for alternate assessments as well. Some states reported that the development of alternate assessments for students with disabilities significantly contributed to the overall increase in state assessment expenditures since the enactment of the NCLB. For example, officials in the four states visited by GAO reported that a general state assessment cost approximately \$30 per student and an alternate assessment cost between \$300 and \$400 per student. Since a small number of students are eligible to take the alternate assessment, the overall impact of assessment expenditures on alternate assessments for students with disabilities as a proportion of the total assessment expenditures remains unclear.

States reported using primarily multiple choice items on assessments, in part to control costs. Thirty-eight of the 48 states that responded to this survey item reported that multiple choice items comprised all or most of the reading assessment, and 39 states reported that multiple choice items comprised all or most of the mathematics assessment. In addition to the cost savings, some states reported that the use of multiple choice items allowed them to score assessments within “challenging” time frames. Since enactment of the NCLB, states have faced pressure to score assessments and report the results prior to the following school year. Some states reported that they considered using more open-ended items; however, it was determined that it was not feasible to score these assessment items within the time period allotted.

State Variability in Expenditures

Previous reports of state expenditures on assessment activities required by the NCLB have clearly demonstrated the variability across states. These reports, however, did not seek to explain the reasons for this variability. A CRS analysis was conducted to determine whether there is evidence

that the variability in state assessment expenditures is due, in part, to differences in state population. This was undertaken because questions have arisen regarding optimal ways to target federal support for state assessments. The current allocation strategy is heavily based on student population.⁶²

Data Source and Questions for Analysis

Survey data were obtained from GAO and used to conduct an analysis of the relationship between state population and expenditures on assessment.⁶³ Data used for this analysis had a number of limitations. First, expenditure data are difficult to collect due to the variability inherent in state and LEA accounting methods.⁶⁴ Second, the expenditure data were collected via a survey of state officials. It is likely that state officials had varying degrees of accuracy in completing the survey, making a comparative analysis less precise. A visual inspection of the data indicated that some state officials may not have understood how to complete certain survey items.⁶⁵ A third limitation of the data was that the responses to survey items could not be linked to individual states. As such, it was not possible to explore other data sources and link them to the primary data source.

Forty-nine of 51 states responded to the survey.⁶⁶ Some respondents did not submit complete survey data. As such, some of the analyses reported here include fewer than 49 cases.⁶⁷ Because data provided by GAO did not include state names or populations, a proxy variable for “state size” was created to conduct the analysis. The state size variable is based on the number of students who took state assessments in mathematics in school year 2007-2008.⁶⁸ States were separated into quartiles by the state size variable.⁶⁹ The quartiles were used to explore the effect of a state’s population on the following outcomes of interest:

1. The percentage of states’ total funding for NCLB assessments provided by the Grants for State Assessments program (Section 6111) and the percentage of states’ total funding for NCLB assessments provided by state government funds (**Table 2**);
2. The total full-time professional equivalent (FTE) state-level expenditures and assessment vendor expenditures for NCLB assessments (**Table 3**); and
3. The ratio of assessment vendor expenditures to state FTE expenditures (**Table 4**).

⁶² For a complete discussion of the current allocation strategy, see the **Appendix**.

⁶³ GAO removed state-identifying information; therefore, state-level conclusions cannot be drawn from this analysis. CRS made decisions regarding the inclusion of incomplete or clearly incorrectly reported data in the GAO survey database; as such, the results of the CRS analysis may differ from those of GAO.

⁶⁴ See “State Expenditures on Assessment Systems Under the NCLB” for an example of how accounting practices may affect the reporting of assessment expenditures.

⁶⁵ For example, some states reported spending relatively small amounts on costly activities (e.g., \$71 for assessment vendors). In several cases, data that were clearly incorrect were removed for the purpose of this analysis.

⁶⁶ For the purpose of this analysis, the District of Columbia is a state.

⁶⁷ The number of states used in each analysis is noted in the data tables.

⁶⁸ All data reported in this section correspond to school year 2007-2008. The number of students who took state assessments in mathematics was chosen over the number of students who took state assessments in reading because more English language learners would have been included in the mathematics assessment, thus representing a more accurate approximation of “state size.”

⁶⁹ States were separated into quartiles at the beginning of the analysis. Because all states did not complete all survey items, the quartiles sometimes contain an uneven number of states.

Results

The analysis highlighted the relationship between federal and state funding of assessments, as well as the relationship between state size and state assessment expenditures. On average, federal formula funding accounts for 43% of states' total expenditures on assessments required by the NCLB (see **Table 2**). The percentage of total state assessment expenditures comprised by federal formula funding is somewhat dependent on state size, with smaller states reporting a larger federal contribution than larger states. The relationship between state size and state government funding of assessments is reversed; that is, smaller states report a smaller contribution from state government funds than larger states.

Expenditures on both state-level staff and assessment vendors have a positive relationship with state size; as state size increases, both kinds of expenditures increase (see **Table 3**). On average, states report spending approximately 16 times more on assessment vendors than on state-level staff dedicated to NCLB assessment activities (see **Table 4**). The analysis is not intended to suggest, however, that hiring more state-level staff to increase capacity would necessarily decrease the cost of assessment vendor services.

Across states with varying populations, consistent average trends emerge. These trends suggest that the population of a state does, to some extent, contribute to the overall cost of state assessment systems. The sizeable variability among states of similar populations, however, suggests that there are factors other than population that are contributing, sometimes significantly, to the cost of assessment systems. Factors that likely contribute to the cost of state assessment systems are design elements of assessment development, scoring, and differences in policy choices and implementation practices (see "Factors that Influence the Cost of State Assessment Systems"). Exploring the relationship between state assessment expenditures and these other factors is not possible with the current data.

Table 2. Percentages of States' Funding for NCLB Assessments from Federal and State Government Funds, by State Size, 2007-2008

	% NCLB Sec. 6111^a	Min^b	Max^c	n^d	% State Government^e	Min^b	Max^c	n^d
All States ^f	43	13	100	45	51	0	84	45
Q1 ^g	65	30	100	10	27	8	70	10
Q2 ^h	39	13	78	11	49	0	83	11
Q3 ⁱ	39	16	80	11	59	20	84	11
Q4 ^j	33	18	59	13	63	34	82	13

Source: CRS analysis of data provided by U.S. Government Accountability Office (GAO), "Survey of States Regarding Student Assessments Under No Child Left Behind Act."

Notes: The percentage of states' funding for NCLB assessments for federal and state government funds may not sum to 100% because additional funds contribute to overall expenditures (e.g., local funds, private funds, nonprofit funds).

- a. The average percentage of states' total reported funding for NCLB assessments from the Grants for State Assessments program in school year 2007-2008 (Section 6111).
- b. Minimum value reported.
- c. Maximum value reported.

- d. Number of states responding to survey item.
- e. The average percentage of states' total reported funding for NCLB assessments from state government funds in school year 2007-2008.
- f. "All states" represents states that responded to specific elements in the GAO survey.
- g. "Q1" is "quartile 1," which represents states that assessed less than or equal to 142,974 students.
- h. "Q2" is "quartile 2," which represents states that assessed more than 142,974 students but less than or equal to 361,845 students.
- i. "Q3" is "quartile 3," which represents states that assessed more than 361,845 students but less than or equal to 627,990 students.
- j. "Q4" is "quartile 4," which represents states that assessed more than 627,990 students.

Table 3. State FTE Expenditures and Assessment Vendor Expenditures for NCLB Assessments, by State Size, 2007-2008

	State FTE Expenditures ^a	Min ^b	Max ^c	n ^d	Assessment Vendor Expenditures ^e	Min ^b	Max ^c	n ^d
All States ^f	\$1,535,665	\$80,000	\$3,900,000	23	\$16,300,472	\$0	\$82,823,000	39
Q1 ^g	\$1,108,226	\$80,000	\$3,900,000	6	\$4,105,709	\$0	\$8,500,000	9
Q2 ^h	\$1,209,292	\$339,980	\$2,975,235	5	\$10,413,185	\$500,000	\$17,839,722	10
Q3 ⁱ	\$1,678,090	\$102,000	\$3,500,000	6	\$17,597,568	\$5,927,533	\$41,200,000	9
Q4 ^j	\$2,092,658	\$890,734	\$3,676,860	6	\$30,568,826	\$1,200,000	\$82,823,000	11

Source: CRS analysis of data provided by U.S. Government Accountability Office (GAO), "Survey of States Regarding Student Assessments Under No Child Left Behind Act."

- a. The average total state FTE cost for NCLB assessments in school year 2007-2008.
- b. Minimum value reported.
- c. Maximum value reported.
- d. Number of states responding to survey item.
- e. States' average expenditures for NCLB assessments that were paid to assessment vendors, contractors, or consultants to develop, administer, score, and report NCLB assessments in school year 2007-2008.
- f. "All states" represents states that responded to specific elements in the GAO survey.
- g. "Q1" is "quartile 1," which represents states that assessed less than or equal to 142,974 students.
- h. "Q2" is "quartile 2," which represents states that assessed more than 142,974 students but less than or equal to 361,845 students.
- i. "Q3" is "quartile 3," which represents states that assessed more than 361,845 students but less than or equal to 627,990 students.
- j. "Q4" is "quartile 4," which represents states that assessed more than 627,990 students.

Table 4. Ratio of Assessment Vendor Expenditures to State FTE Expenditures, by State Size, 2007-2008

	Ratio of Assessment Vendor to State FTE Expenditures	Min ^a	Max ^b	n ^c
All States ^d	15.67	0.13	57.90	20
Q1 ^e	14.66	0.13	43.24	6
Q2 ^f	9.92	1.47	16.07	5
Q3 ^g	12.34	2.23	27.18	4
Q4 ^h	25.28	10.17	57.90	5

Source: CRS analysis of data provided by U.S. Government Accountability Office (GAO), “Survey of States Regarding Student Assessments Under No Child Left Behind Act.”

- a. Minimum value reported.
- b. Maximum value reported.
- c. Number of states responding to survey item.
- d. “All states” represents states that responded to specific elements in the GAO survey.
- e. “Q1” is “quartile 1,” which represents states that assessed less than or equal to 142,974 students.
- f. “Q2” is “quartile 2,” which represents states that assessed more than 142,974 students but less than or equal to 361,845 students.
- g. “Q3” is “quartile 3,” which represents states that assessed more than 361,845 students but less than or equal to 627,990 students.
- h. “Q4” is “quartile 4,” which represents states that assessed more than 627,990 students.

Promoting Efficiency in State Assessment Systems

Moving forward, Congress is likely to grapple with decisions regarding how much support to provide for state assessments, how to allocate funds, and whether any practices can be supported to encourage efficiency. As has been noted, assessment practices and the assessment industry have changed dramatically following the passage of the NCLB, and there is a dearth of current research on matters related to the cost of varied assessment practices. Based on the literature that is available, and insights garnered through discussions with assessment vendors, this section discusses options for containing costs and possible tradeoffs associated with them.

As Congress considers the reauthorization of the ESEA, what will the federal government’s role be in funding state assessments? Should federal funding aim to fund a certain proportion of state assessments required by federal law? Given the flexibility that states have to design their own assessment systems, it may be difficult for federal funding to meet a constantly moving target. The federal government could, however, develop policies that promote quality and efficiency in state assessment systems. Federal policies that aim to promote efficiencies, however, may risk reducing the quality of assessments if the efficiencies are supported without full consideration of the consequences. Certain efficiencies may involve tradeoffs in terms of the type of assessments, number of assessments, and timeline imposed on states to report results. Efficiencies may also require balancing state independence with promoting state consortia in order to realize economies of scale. The remainder of this report examines the potential advantages and consequences of these tradeoffs and compromises in the design of state assessment systems.

Some educators have discussed dissatisfaction with the quality of state assessment systems and the over-reliance on multiple choice items. One proposed alternative is to increase the use of performance-based assessment in state assessment systems. A performance-based assessment is an authentic task that assesses what a student knows and can do. It can take many forms, including conducting science assessments, writing persuasive essays, completing research projects, or applying mathematical problem solving to a real-world scenario.⁷⁰ A recent study analyzed the costs involved for states to move from primarily multiple choice assessments to a “high-quality assessment” (HQA)⁷¹ system, which would utilize more performance-based items.⁷² The researchers found that the cost of a new HQA system would be approximately three times higher than a traditional multiple choice assessment. Implementing the HQA system with certain cost reduction strategies, however, led to a significant decrease in cost.⁷³ In a hypothetical scenario where all cost reduction strategies were used, the HQA system was projected to be less expensive than traditional multiple choice assessments.⁷⁴

The cost reduction strategies used in the aforementioned study are not unique to a specific type of assessment. These strategies can be employed with both traditional and innovative assessment systems. The implications of using certain cost reduction strategies to promote efficiency in assessment systems are discussed below.

Scoring Practices

As discussed earlier, most state assessment systems use multiple choice assessments or multiple choice assessments with limited constructed response items. Currently, it is most common for states to use automated scoring for multiple choice items and human scoring for constructed response items. As part of their contract, assessment vendors are typically responsible for the human scoring of constructed response items, which represents one of the most costly and time-consuming activities in a state assessment system. In order to move away from multiple choice assessments, states would need to find a way to reduce the cost of human scoring.

Several options have been proposed to reduce scoring costs. One option is the use of distributed scoring. Distributed scoring involves scorers working from home, accessing the assessment items through an online platform. In this arrangement, the assessment vendor is not required to supply a facility or equipment for scorers. In reality, it is likely that assessment vendors would use a mixture of onsite and distributed scoring. When investigating cost strategies, researchers modeled

⁷⁰ Some researchers make a distinction between a “performance event” (PE) and a “performance task” (PT). A PE is defined as an on-demand activity that students complete in a class period in school. A PT is defined as an activity that students will work on in class and outside of class for periods ranging from a couple of days to several weeks. For the purpose of this report, both PE and PT are considered under the term “performance-based assessment.”

⁷¹ The term “high-quality assessment” is used by the authors of the study. It is not possible for CRS to determine if the assessment in this study is of a high quality; nevertheless, the terms “high-quality assessment” and “HQA” are used in this report in an effort to maintain consistent terminology with the study.

⁷² Barry Topol, John Olson, and Ed Roeber, *The Cost of New Higher Quality Assessments: A Comprehensive Analysis of the Potential Costs for Future State Assessments*, Stanford Center for Opportunity Policy in Education, Stanford, CA, 2010, http://edpolicy.stanford.edu/pages/pubs/pub_docs/assessment/scope_pa_topol.pdf.

⁷³ The cost reduction strategies investigated in this study included various scoring options, online test delivery, and the use of state consortia.

⁷⁴ See footnote 72.

the savings involved in using a 50-50 mix of onsite and distributed scoring. The 50-50 mix resulted in an 11% reduction in the costs of scoring constructed response items.⁷⁵

Another option that has been proposed to reduce scoring costs is using teachers to score assessments. Teachers evaluate student performance on a daily basis, and it may be a natural extension of their job to score state assessments. Furthermore, teachers would be able to score the full range of assessment types, including multiple choice, constructed response, and performance-based assessments. Teacher scoring could be implemented in several ways. For example, teachers could score assessments as part of their expected duties or as a professional development activity. Alternatively, teachers could be paid a stipend for the extra time involved with training and scoring assessments.

A number of issues may arise, however, if teachers score assessments. First, since teachers have a stake in the outcome of the NCLB accountability system, the state would need procedures in place to ensure objective scoring. In the NCLB accountability system, states must hold schools and LEAs accountable for student performance; therefore, it may not make sense to have teachers score assessments within their school or LEA, given that they have a vested interest in the outcome. It may be possible to require teachers to score assessments from other schools or other LEAs; however, the potential travel involved may negate the cost saving potential of using teachers.⁷⁶

Another option to consider in reducing scoring costs is the advent of artificial intelligence (AI). AI software has been developed to score some constructed response items.⁷⁷ Some research finds that AI produces results that are relatively comparable to human scoring;⁷⁸ however, the cost for AI scoring is currently too high to realize any savings in a state assessment system. There is speculation that the costs of AI scoring will come down considerably in the next several years, and it may serve as a cost reduction strategy in the future.⁷⁹

A related issue to consider is the time required to score assessments. Under the NCLB, the state assessment is typically administered in the spring and results must be reported before the beginning of the next school year. The timeline is driven by regulations that require notifying parents of their public school choice options under the NCLB, which is dependent upon the results of state assessments. Regulations specify that an LEA must notify parents of public school choice options made available to students who attend a school in need of improvement. Parental notification “must be made sufficiently in advance of, but no later than 14 calendar days before, the start of the school year so that parents have adequate time to exercise their choice option before the school year begins.”⁸⁰ Since decisions regarding the determination of schools in need

⁷⁵ See footnote 72.

⁷⁶ Travel expenditures may be an issue in cases where teachers are evaluating performance-based assessments that require observation.

⁷⁷ For example, AI could be used to score constructed response items that require a written response to a prompt; however, it is not yet capable of scoring performance-based assessments.

⁷⁸ The Educational Testing Service reports a 0.9 correlation between AI and human scoring. Mike Bowler, “A Scary Future: MSPAP Online; Technology: Electronic Services are Vying for the Business of Creating and Scoring Standardized Tests,” *The Sun*, April 24, 2002, p. 2.B, Final Edition.

⁷⁹ See footnote 72 and http://www.ets.org/research/capabilities/automated_scoring.

⁸⁰ 34 C.F.R. §200.37.

of improvement affect public school choice options, assessment results must be available no later than 14 calendar days before the start of the school year.

If multiple choice assessments were to be replaced by assessments with more constructed response items or performance items, there would be an increase in the amount of human scoring. To accommodate the increase in human scoring, states may need an extended window of time to score assessments. This extended window of time may require either (1) testing students earlier in the year, or (2) allowing results to be reported after the start of the next school year. If students are tested earlier in the year, the results of the assessment may not accurately represent a full year of instruction, and the results may not reflect a valid measurement of grade-level academic content and performance standards. On the other hand, if results were allowed to be reported later (i.e., after the start of the next school year), parents would not have the ability to exercise the public school choice option included in the NCLB accountability system.

State Consortia

Under the NCLB, states were afforded the flexibility to develop their own state academic content and performance standards and to develop state assessments that measure student achievement against those standards. This flexibility created the possibility that over 50 unique state assessment systems would emerge. Each assessment system had the potential to differ in terms of assessment type, assessment development, scoring procedures, administration procedures, timing of the assessment, and so forth. Because each state was focused on developing assessments that would measure specific state standards, there was limited interest in pooling resources across states to develop common assessments.

As discussed earlier, there is currently one state consortium that works to develop, maintain, and administer assessments that fulfill the requirements of the NCLB.⁸¹ Due to the reported success of this consortium, interest in the use of state consortia has increased. The use of state consortia may promote efficiency in state assessment systems by allowing states to share the development and overhead costs. One study modeled the potential cost savings of state consortia comprised of 10, 20, and 30 states. Participating in a 10-state consortium was estimated to reduce average per pupil costs on assessments by 24% compared to the costs of a single state. Participation in a 20-state consortium reduced average per pupil costs by 27% compared to a single state, and participation in a 30-state consortium reduced average per pupil costs by 30% compared to a single state.⁸² Although states would continue to realize savings with larger consortia, the cost model indicates that the majority of the savings are realized in a 10-state consortium. If states choose to enter a consortium in the future, they may consider the incremental savings of a larger consortium versus the ease of working with a smaller consortium.

⁸¹ See the previous section of this report, “Policy Choices and Implementation Practices.”

⁸² The percentages reported here correspond to savings that consortia of states would realize when implementing a high-quality assessment system as described in the study. The findings cannot necessarily be generalized to current assessment systems used by states. For the purpose of this report, the relative savings between 10-, 20-, and 30-state consortia is relevant to the discussion. To view the parameters of the high-quality assessment system that is used in this cost analysis, see Barry Topol, John Olson, and Ed Roeber, *The Cost of New Higher Quality Assessments: A Comprehensive Analysis of the Potential Costs for Future State Assessments*, Stanford Center for Opportunity Policy in Education, Stanford, CA, 2010, http://edpolicy.stanford.edu/pages/pubs/pub_docs/assessment/scope_pa_topol.pdf.

ED has taken steps to promote the use of state consortia through its Race to the Top Assessment Program (RTTT Assessment Program). The RTTT Assessment Program is a \$350 million program focused specifically on funding comprehensive assessment systems for state consortia.⁸³ The Secretary announced the winners of the RTTT Assessment Program on September 2, 2010.⁸⁴ Two state consortia received awards under this program: (1) Partnership for Assessment of Readiness of College and Careers (PARCC; 26 states), and (2) SMARTER Balanced Assessment Consortium (SBAC; 31 states).⁸⁵ Across the two applications, 45 states are participating in at least one of the consortia.⁸⁶

Both PARCC and SBAC proposed to develop assessments that measure student knowledge and skills across a common set of college- and career-ready standards in mathematics and English language arts for grades 3 through 8 and one grade in high school.⁸⁷ The assessments must assess all students, including English language learners and students with disabilities.⁸⁸ Because the RTTT Assessment Program grants were awarded relatively recently, it is unclear how they may affect state assessments used for accountability purposes. ED's "expectation" is that states adopting assessment systems developed through this program will use them to meet the requirements for state assessment systems under Title I of the NCLB.⁸⁹ The RTTT Assessment Program, however, is a one-time competition, and it is unclear whether any continued support for participating state consortia would be considered. Ultimately, decisions regarding the type of state assessments that meet the requirements of Title I may be determined by Congress during the next reauthorization of the ESEA.

Technology

As technology becomes less expensive, online delivery of state assessments may lower the cost of state assessment systems. Using online delivery may potentially reduce costs associated with production, manufacturing, and shipping assessments from the vendor to schools. Some states, such as Oregon and Virginia, have already moved away from paper-and-pencil tests to online delivery, and it is expected that other states may follow.⁹⁰ In one study, the cost of online delivery

⁸³ Authority to administer the RTTT Assessment Program was included in the American Recovery and Reinvestment Act (P.L. 111-5).

⁸⁴ See <http://www.ed.gov/news/press-releases/us-secretary-education-duncan-announces-winners-competition-improve-student-asse>.

⁸⁵ Several states participate as non-governing member states in both consortia. PARCC was awarded approximately \$170 million; SBAC was awarded approximately \$160 million.

⁸⁶ For the purpose of this report, the District of Columbia is considered to be a participating state.

⁸⁷ The proposed assessments must be able to measure student achievement and student growth.

⁸⁸ For the purpose of the RTTT Assessment Program, a student with a disability is defined as "a student with a disability under the Individuals with Disabilities Education Act, as amended (IDEA), except for a student with a disability who is eligible to participate in alternate assessments based on alternate achievement standards consistent with 34 CFR 200.6(a)(2)". For information on alternate assessments based on alternate achievement standards, see CRS Report R40701, *Alternate Assessments for Students with Disabilities*, by (name redacted).

⁸⁹ ED expresses this expectation in the *Guidance and FAQs* document on pages 4 and 24.

⁹⁰ Loren Moreno, "Key Test for No Child Left Behind Going Online in 2011," *Honolulu Advertiser*, September 3, 2009; Michael Alison Chandler and Amit R Paley, "Online Glitches Frustrate Test-Takers in VA Schools," *The Washington Post*, May 17, 2007, p. B.1, Final Edition; Barry Topol, John Olson, and Ed Roeber, *The Cost of New Higher Quality Assessments: A Comprehensive Analysis of the Potential Costs for Future State Assessments*, Stanford Center for Opportunity Policy in Education, Stanford, CA, 2010, p. 11, http://edpolicy.stanford.edu/pages/pubs/pub_docs/assessment/scope_pa_topol.pdf.

was about 12% less than paper-and-pencil tests; however, the use of technology was combined with other cost reduction strategies, which confounds the actual savings of online delivery. Furthermore, the study did not include the additional cost of purchasing equipment in its estimate, since this cost is highly variable depending on the number of computers already available.⁹¹

One of the major challenges of online testing is ensuring that all schools have the necessary equipment to allow for large-scale assessment in a relatively short period of time. A school's readiness to implement online testing may be a reflection of the resources available at the school, which may inadvertently highlight the technology gap between schools. Furthermore, it is not known at this time how often schools would need to replace equipment in order to work with new software or online delivery mechanisms that are constantly changing. Federal, state, and local resources may need to provide ongoing support for the purchase of equipment if states are to develop and maintain online delivery systems for state assessments.

Another concern with the use of online delivery of assessments is the potential for technological problems that may lead to delays in testing or scoring. In the worst case scenario, results could be lost completely. Some technological glitches have already occurred in Virginia. In some cases, students could not complete their online test due to a glitch in the server maintained by the assessment vendor. The students were required to retake the assessment on a different day, which raises some concern about the validity of the scores from the second administration.⁹² If online delivery moves forward, states may need to ensure that schools have the infrastructure in place to carry out large-scale assessment and develop policies to deal with technological glitches in the administration or scoring of state assessments.

⁹¹ See footnote 72.

⁹² Michael Alison Chandler and Amit R Paley, "Online Glitches Frustrate Test-Takers in VA Schools," *The Washington Post*, May 17, 2007, p. B.1, Final Edition.

Appendix. Federal Funding for State Assessments

To assist states in complying with the new assessment requirements, the NCLB authorized two programs that specifically provide funding for the development and administration of state assessments: Grants for State Assessments (Section 6111) and Grants for Enhanced Assessment Instruments (Section 6112). To carry out these programs, \$490 million was authorized to be appropriated in FY2002, and such sums as may be necessary were authorized for each of the five succeeding fiscal years.⁹³ In general, funds up to the “trigger amount” (discussed below) are distributed by formula grants under the Grants for State Assessments program and the remaining funds are distributed through a competitive process under the Grants for Enhanced Assessment Instruments.⁹⁴ The remainder of this section discusses the formula grant program Grants for State Assessments.

Under the Grants for State Assessments program, statutory provisions specify that the Secretary shall make grants to states for the following purposes:

- to pay the costs of the development of the additional state assessments and standards required by Section 1111(b) of the NCLB, which may include the costs of working voluntary partnerships with other states, at the sole discretion of each state; and
- if a state has developed the assessments and standards required by Section 1111(b) of the NCLB, to administer assessments or carry out other activities related to ensuring that the state’s schools and LEAs are held accountable for the results.

The assessment requirements under Title I-A of the NCLB are contingent upon the appropriation of minimum annual amounts for state assessment grants; this minimum annual amount is referred to as the “trigger amount” in the formula allocation. The “trigger amount” is defined in legislation as the amount made available to carry out Sections 6111 and 6112 minus an amount specified in Section 1111(b)(3)(D).⁹⁵ For each of FY2002-FY2008, at least the minimum “trigger amounts” have been appropriated for these grants. The authorization of appropriations for the assessment grant programs expired in FY2008; however, these programs have continued to receive an appropriation above the “trigger amount” specified for the most recent year of authorization (i.e., \$400 million in FY2008).

From the “trigger amount,” Grants for State Assessments are allocated as follows:

1. The Secretary shall reserve 0.5% of the total for the outlying areas and 0.5% of the total for the Bureau of Indian Affairs (BIA).

⁹³ Although this program is no longer authorized, Congress has continued to provide appropriations.

⁹⁴ For more information on this program, see CRS Report R40514, *Assessment in Elementary and Secondary Education: A Primer*, by (name redacted).

⁹⁵ If the minimum “trigger amount” was not appropriated, the administration (but not the development) of grades 3-8 and science assessments may have been delayed by one year for each year that the “trigger amount” was not met. “Trigger amounts” for each year since the enactment of NCLB were \$370 million in FY2002, \$380 million in FY2003, \$390 million in FY2004, and \$400 million in FY2005-FY2008.

2. From the remainder of the total, the Secretary shall allocate to each state an amount equal to \$3 million.
3. From the remainder of the total, the Secretary shall allocate funds among states in proportion to their number of children and youth aged 5-17 years.

All states, the District of Columbia, and Puerto Rico receive funding under the Grants for State Assessments program. **Table A-1** shows state appropriations for the Grants for State Assessment program for FY2002-FY2011 (estimate).

**Table A-1. Federal Assessment Grants to States Under Section 6111:
FY2002-FY2011 (estimate)**

(In dollars)

State	FY2002- FY2008	FY2009	FY2010	FY2011 (estimate)	Total
Alabama	45,400,442	6,627,669	6,617,666	6,617,666	65,263,443
Alaska	25,139,278	3,571,366	3,573,928	3,573,928	35,858,500
Arizona	53,529,242	8,326,381	8,395,851	8,395,851	78,647,325
Arkansas	35,902,804	5,237,350	5,246,143	5,246,143	51,632,440
California	226,508,317	32,776,448	32,708,959	32,708,959	324,702,683
Colorado	45,818,050	6,795,252	6,838,322	6,838,322	66,289,946
Connecticut	39,599,364	5,685,190	5,656,669	5,656,669	56,597,892
Delaware	25,305,379	3,656,838	3,654,778	3,654,778	36,271,773
District of Columbia	23,338,943	3,338,296	3,341,877	3,341,877	33,360,993
Florida	106,414,081	15,803,919	15,857,370	15,857,370	153,932,740
Georgia	71,002,744	11,085,043	11,147,998	11,147,998	104,383,783
Hawaii	27,327,545	3,885,424	3,895,349	3,895,349	39,003,667
Idaho	29,244,147	4,300,594	4,308,311	4,308,311	42,161,363
Illinois	91,364,303	13,215,809	13,152,557	13,152,557	130,885,226
Indiana	55,789,075	8,104,086	8,085,387	8,085,387	80,063,935
Iowa	36,567,200	5,285,999	5,263,817	5,263,817	52,380,833
Kansas	36,188,997	5,226,374	5,221,169	5,221,169	51,857,709
Kentucky	42,683,954	6,234,621	6,229,277	6,229,277	61,377,129
Louisiana	46,290,765	6,564,556	6,574,511	6,574,511	66,004,343
Maine	27,546,131	3,909,442	3,891,080	3,891,080	39,237,733
Maryland	51,165,837	7,331,511	7,318,720	7,318,720	73,134,788
Massachusetts	53,350,583	7,665,358	7,655,840	7,655,840	76,327,621
Michigan	77,395,049	10,889,892	10,708,981	10,708,981	109,702,903
Minnesota	48,727,313	7,006,812	6,987,447	6,987,447	69,709,019
Mississippi	37,599,325	5,440,764	5,424,240	5,424,240	53,888,569
Missouri	52,018,829	7,569,472	7,569,841	7,569,841	74,727,983

State	FY2002- FY2008	FY2009	FY2010	FY2011 (estimate)	Total
Montana	25,896,247	3,711,984	3,699,787	3,699,787	37,007,805
Nebraska	30,617,643	4,407,939	4,409,107	4,409,107	43,843,796
Nevada	33,636,983	5,095,238	5,122,895	5,122,895	48,978,011
New Hampshire	27,911,709	3,974,956	3,953,185	3,953,185	39,793,035
New Jersey	67,321,712	9,662,547	9,627,350	9,627,350	96,238,959
New Mexico	31,987,479	4,583,311	4,592,850	4,592,850	45,756,490
New York	121,180,305	17,305,137	17,232,290	17,232,290	172,950,022
North Carolina	66,131,848	10,112,747	10,172,294	10,172,294	96,589,183
North Dakota	24,280,027	3,452,253	3,447,154	3,447,154	34,626,588
Ohio	83,086,800	11,882,258	11,780,417	11,780,417	118,529,892
Oklahoma	\$40,106,191	5,859,167	5,876,636	5,876,636	57,718,630
Oregon	39,773,865	5,790,331	5,778,437	5,778,437	57,121,070
Pennsylvania	84,784,653	12,051,777	12,018,306	12,018,306	120,873,042
Puerto Rico	44,399,566	6,321,855	6,246,417	6,246,417	63,214,255
Rhode Island	26,409,085	3,749,371	3,741,119	3,741,119	37,640,694
South Carolina	43,497,723	6,412,299	6,422,322	6,422,322	62,754,666
South Dakota	25,305,350	3,624,794	3,622,361	3,622,361	36,174,866
Tennessee	51,789,680	7,749,391	7,747,147	7,747,147	75,033,365
Texas	154,131,874	24,006,948	24,441,590	24,441,590	227,022,002
Utah	36,667,615	5,596,409	5,645,936	5,645,936	53,555,896
Vermont	24,168,279	3,430,537	3,417,058	3,417,058	34,432,932
Virginia	59,951,570	8,814,697	8,842,476	8,842,476	86,451,219
Washington	54,334,554	7,954,146	7,975,183	7,975,183	78,239,066
West Virginia	29,673,503	4,255,119	4,247,038	4,247,038	42,422,698
Wisconsin	50,617,806	7,257,019	7,206,912	7,206,912	72,288,649
Wyoming	23,720,236	3,403,304	3,407,645	3,407,645	33,938,830
All States	2,712,600,000	396,000,000	396,000,000	396,000,000	3,900,600,000

Source: Data reported from the U.S. Department of Education. See James Taylor, Brian Stecher, and Jennifer O'Day, et al., *State and Local Implementation of the No Child Left Behind Act*, U.S. Department of Education (ED), Volume IX—Accountability Under NCLB: Final Report, Washington, DC, 2010, <http://www2.ed.gov/rschstat/eval/disadv/nclb-accountability/nclb-accountability-final.pdf>.

Author Contact Information

(name redacted)
Analyst in Education Policy
/redacted/@crs.loc.gov, 7-....

Acknowledgments

CRS Graphics Specialist Amber Wilhelm created the graphics in 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.