

# Implementing the No Child Left Behind Act: Challenges for the States

---

*Margaret E. Goertz*  
*Graduate School of Education*  
*University of Pennsylvania*

The No Child Left Behind Act of 2001 (NCLB) requires states to test specified subjects and grades, to establish minimum performance standards for students, schools and school districts, and to provide assistance and impose sanctions on schools and districts that do not meet performance goals as a condition of receiving federal aid. NCLB builds on earlier federal and state education reform initiatives, and its success depends on the willingness and capacity of states and localities to enact its more stringent provisions. This article examines challenges facing states as they implement the new law and concludes that most lack the human and fiscal resources to support school improvement.

Federal policy has played a major role in supporting standards-based reform since the passage of the Improving America's Schools Act (IASA) of 1994. That law required states to establish challenging content and performance standards, implement assessments that measure students' performance against these standards, hold schools and school systems accountable for the achievement of all students, and take other steps to promote programmatic flexibility and to foster instructional and curricular reform.

During the mid- and late 1990s, states and school districts began to move in the direction of standards-based reform, consistent with the intent of IASA, but state policy responses were uneven. Although all states developed assessments, standards, performance reporting, and, in most cases, consequences for performance, states found different ways to define what

I thank Sarah Costelloe for her assistance in collecting data for this article.  
Requests for reprints should be sent to Margaret E. Goertz at [cpre@gse.upenn.edu](mailto:cpre@gse.upenn.edu)

it meant for schools to succeed, what indicators to include in their definition of success, and what the consequences would be (Goertz & Duffy, 2001). The No Child Left Behind Act of 2001 (NCLB) was designed, in part, to address this variability in state policy. With the enactment of this law, the federal government expanded its role significantly, requiring states to test more and set more ambitious and uniform improvement goals for their schools, and prescribing sanctions for schools that fail to meet these goals.

Yet, as with IASA, successful implementation of NCLB will depend on the willingness and capacity of states and localities to enact policies and practices that reflect both the spirit and the letter of the law. There appears to be widespread support at the state and school district level for the basic premises of NCLB—that having a uniform accountability system based on content and performance standards and focusing attention on subgroup performance and achievement gaps will positively affect student achievement (Center on Education Policy [CEP], 2004). At the same time, the Act has generated criticism from educators, state and local policymakers, and some members of Congress. Legislatures in Ohio, Virginia, and Utah have questioned the “unfunded mandates” that have been created by the NCLB legislation. Senator Edward Kennedy and other Democratic congressmen argue that federal funding is insufficient to meet the demands of the law. Other issues concern the inclusion of students with disabilities and English language learners (ELLs) in state assessment and accountability systems, the large number of schools identified as not making adequate yearly progress, the cost of expanding assessments and providing student supports, the lack of alignment between federal and state accountability designs, and the difficulty of applying the “highly qualified teacher” criteria to special education and middle school teachers and to teachers in rural school districts.

The U.S. Department of Education has taken steps to address some of these concerns. For example, guidance issued in early 2004 gave states a little more flexibility in how they include special needs students in their accountability calculations and in how they demonstrate that certain categories of teachers are “highly qualified.” However, from the state and local perspective, many issues remain unresolved. This article looks at implementation challenges in four of these areas: assessment, accountability, special needs students, and the human and fiscal capacity to support school improvement.

### Assessment

The NCLB Act expands testing in reading and mathematics significantly, from at least once during each of three grade spans—3rd to 5th, 6th to 9th, and 10th to 12th—to every student, annually, in Grades 3 through 8,

and once in high school. In addition, the law requires states to assess students in science once during each of the three grade spans starting in 2007–08, and requires local school districts to assess the English proficiency of ELLs starting in 2002–03. At the time of the law’s enactment, 48 states had implemented statewide assessments in reading and mathematics. However, only 13 states and the District of Columbia tested consecutive grades between Grades 2 or 3 and at least Grade 8 in the same subject areas using the same assessment, as required by the NCLB. Another 3 states tested consecutive grades between Grades 2 or 3 and 8 in different subjects and/or using multiple assessments. The other 32 states tested students in only one or two grades per subject in elementary school, middle school, and high school, as required by the IASA. Thirty-four states tested science, although not always in all of the three grade spans (Goertz & Duffy, 2001; NATI/IRP, 2001; U.S. Department of Education, 2001). Thus, most states must expand the size and scope of their assessment programs significantly, adding anywhere from 1 to 11 new tests (General Accounting Office [GAO], 2003).

This expansion has major cost and capacity consequences for states. National estimates of the cost of developing and administering new state assessment programs range from \$1.9 billion to \$7 billion, depending on the methodology used to make the cost estimates and assumptions about the mix of items on the assessments (multiple choice, open ended, essay), public release of test items, and the extent of test customization (GAO, 2003). The federal government authorized \$2.34 billion between fiscal year (FY) 2002 and FY2007 to support state assessment programs. States must absorb the additional cost. Ohio, for example, projects that the cost of administering the newly required achievement tests will exceed earmarked federal support by \$9 million a year (Driscoll & Fleeter, 2003). A few states are finding needed funds by eliminating tests in science (as least temporarily), social studies, and other subjects not covered by NCLB, which could have the unintended consequence of narrowing the focus of educators on the tested subjects of reading and mathematics.

The impact of expanded state testing on districts will be mixed. Under IASA, many districts already tested additional grades to measure the continuous progress of students toward district and/or state goals, and to provide instructional feedback to teachers and schools (Goertz, 2001). Therefore, the increased “test burden” on students may be limited. This is particularly true in districts that “back-filled” with tests that are similar to state assessments in type and coverage. Districts that use more performance-based or instructionally based assessments, such as running records or benchmark examinations, face a dilemma, however. They can

continue their assessment programs and test students more, or, facing substantial test burden, they may eliminate their local tests.

### Accountability

When Congress enacted NCLB, states were at different stages of meeting the accountability requirements of the new law. Though the intent of the IASA was to create single and “seamless” accountability systems that would treat both Title I and non-Title I schools alike, only 22 states had single accountability systems in place by 2000–2001. More than half of the states had dual accountability systems where Title I schools were subject to different measures of adequate yearly progress (AYP). Thirteen of these states used public reporting as their primary accountability mechanism for non-Title I schools. States also differed in the percentage of students that schools were expected to bring up to their “proficient” standard, time lines for meeting these performance goals (if they had time lines), and how they measured AYP. Only two states required schools to close achievement gaps between students of different racial-ethnic groups, economic status, disability status, or English language proficiency. No state had a system in place that met all of the provisions of NCLB (Goertz & Duffy, 2001).

There has been considerable discussion in the policy, political, and research communities about the type and extent of flexibility states have in responding to the requirements of the NCLB Act. Regarding accountability, some provisions allow no flexibility. For example, some state accountability systems used a performance index that combined performance on reading and mathematics (a compensatory model). States now have to hold schools and districts accountable separately for reading and mathematics (a conjunctive model). States must calculate AYP separately for all students and all subgroups of students in a school and school district. Also, states must follow the NCLB provisions for establishing “starting points” for the AYP targets, and time lines for implementing specified sanctions to Title I schools. States still establish their own standards in reading, mathematics, and science and design their own assessments, which must be aligned with state standards. Other areas of flexibility include setting the growth trajectory for improvement and the minimum number of students required for inclusion in accountability calculations, and determining whether NCLB sanctions should be applied to non-Title I schools. And states may choose how to incorporate NCLB accountability requirements into their state accountability systems.

All states have developed accountability policies that meet the requirements of NCLB. *Education Week's* 2003 state policy survey shows that about

half of the states will use the NCLB model of AYP as their state accountability system (Olson, 2003). Some, like New Hampshire, did not have a state-wide accountability system (beyond public reporting) prior to the enactment of NCLB. Other states, like New Jersey and Maryland, replaced their prior definition of AYP. Under Maryland's old system, for example, schools had to show "substantial and sustained" progress toward state performance standards as calculated by a School Performance Index to make AYP. Its new accountability plan sets separate and relatively linear performance targets for reading and mathematics, based on a school's starting point and the goal of having 100% students achieve proficiency.

Twenty-three states, however, will include criteria in addition to those required under NCLB provisions in their state accountability systems. At least 11 states will apply two or more ratings to their schools, generally the federal AYP rating and a separate state rating. Some states rate their schools using letter grades (e.g., Florida and Michigan); others describe a continuum of performance that ranges from Excellent to Unsatisfactory, or Excelling to Underperforming. Although the labels differ, these state accountability systems share a set of design features that distinguish them from the provisions of NCLB.

### *Use of Performance Indexes*

Under NCLB, states hold schools accountable separately for reading and mathematics. Many of the state accountability systems, however, continue to use indexes that combine performance on state assessments. Some states average test scores across the subject areas; a few give differential weight to subjects. Some states include subject areas beyond reading and mathematics and some incorporate performance on nonacademic indicators. Kentucky's Accountability Index, for example, includes student performance in seven academic areas, attendance, retention, and dropout rates. Michigan's Education Yes! program includes performance in four subject areas and gives one third of the weight to a number of school climate and organization measures, such as instructional quality and learning opportunities. Kentucky and other states also give differential weight to different levels of student performance. A Kentucky student who performs at the "medium novice" level receives 13 points. Students at the "low apprentice" level receive 40 points; at the "medium apprentice" level, 60 points; at the "proficient" level, 100 points; and at the "distinguished" level, 140 points. Thus, schools have an incentive to move students to increasingly higher levels of performance, particularly at the lowest levels of achievement. In its early review of state NCLB plans, the U.S. Department of Education did not permit states to use these differential weights in their

AYP calculations. It then changed its position but does not allow indexes that allocate additional points to students scoring above the proficient level.

### *Use of Growth and Status Measures*

Many states rate their schools on the extent to which they meet a state-defined growth target. California, for example, expects schools to raise their Academic Performance Index (API)—a composite of student performance on English/language arts, mathematics, and social studies tests—by 5% of the distance between their base for a given year and the statewide API performance target of 800. A school that is ranked in the bottom half of the API distribution and does not meet or exceed its growth target may be identified for state assistance and intervention. Like many states, California measures improvement by the degree to which a school average changes over time. A few states, like North Carolina and South Carolina, track the progress of individual students over time instead. This approach is sometimes called a “value-added” measure of schooling and is being considered or adopted by a growing number of states, especially as they expand their testing systems to cover Grades 3 through 8. North Carolina provides monetary rewards to schools that meet or exceed their growth targets, regardless of overall performance. Schools that do not meet their growth target and who have fewer than half of their students meeting the state performance standard are identified as low-performing and are assigned state assistance teams. Colorado and Massachusetts also hold schools accountable for both absolute performance and academic improvement.

### *Lack of Subgroup Performance*

Most of the state accountability systems do not include subgroup performance in their own rating systems. At least one state views the federal AYP requirements as the “gap-closing” component of its state system, a way of holding schools accountable for closing achievement gaps between different groups of students.

States that have retained their own accountability systems are linking their state and federal AYP measures in different ways. A few states use their state ratings to make federal AYP determinations. California and Kentucky, for example, use their state accountability index as the “other academic indicator” required under NCLB. In California, a school must show growth of at least 1 point on the API or have an API score of at least 560 to make AYP. (This is a different, and less stringent, definition of progress than is used in the state accountability system, however.) Florida

schools with a state grade of D or F cannot be designated as making AYP. A few states incorporate federal AYP into their state accountability measures, adjusting a school's state accountability designation upward (a compensatory model) or downward (a conjunctive model) based on its AYP status. For example, a school that makes AYP can be rated no lower than a C in Michigan or lower than "continuous improvement" in Ohio, regardless of performance on state indicators. On the other hand, no school in Michigan can receive a grade of A if it does not make AYP, and schools rated as excellent and good in South Carolina will have their ratings lowered by one level if they do not meet their schoolwide AYP targets. Arizona includes the schoolwide AYP rating as only one of several factors in determining a school's accountability classification.

### *Effect of State Policies*

These policy design decisions—particularly the use of growth models and lack of subgroup accountability—affect which schools will be identified for improvement, and therefore are eligible for assistance and possibly subject to sanctions under federal and state laws. In many of these states, considerably more schools were identified under NCLB than under their state systems. For example, in Florida, though close to 90% of the state's schools failed to make AYP under the federal guidelines in 2002–03, only 6% of the schools were given grades of D and F. Nearly half of the schools received a grade of A. In North Carolina, 47% of schools made AYP, whereas 90% of the state's schools were deemed as meeting "state growth goals." Similarly, only 55% of California's schools made federal AYP, despite the fact that 78% of the state's schools met the API (Hoff, 2004).

These disparate ratings have created a public relations nightmare for states, particularly in suburban school districts that are highly rated under state measures. Some state policymakers vouch for the validity of their measures over those of NCLB. For example, on releasing the list of schools that failed to make AYP—a number that far exceeded those identified as underperforming by the state—the state superintendent in Arizona urged parents to look at state, as opposed to federal, school rankings. He argued that the state formula is much more complex, taking into consideration factors such as Stanford 9 scores and incremental improvement of a school's overall state test scores. "We worked hard to make the state system fair and accurate, whereas the federal system has a fatal flaw. ... The overriding result is to overidentify schools that fail" (Kossan, 2003). Differing accountability determinations also led the chief state school officers of 14 states to request that NCLB be amended to permit the use of growth models for demonstrating academic gains, even if these gains fall short of what is re-

quired for students to meet NCLB proficiency targets.<sup>1</sup> They argued that NCLB, as currently written, will label the vast majority of schools as in need of improvement and divert limited resources to schools that have shown “steady and significant improvement for all groups of students” (Schemo, 2004). Finally, states are wrestling with how to validate their accountability measures. Are their data accurate? Are they identifying the right schools as in need of improvement?

### Special Needs Students

Probably the most controversial and contentious provisions of the NCLB are those that apply to special needs students, particularly students with disabilities and ELLs. These populations have received heightened attention (as was intended) because the subgroup accountability provisions of NCLB account, in part, for the growing number of schools identified as not making AYP. In 2003–04, nearly one fourth of school districts reported they had at least one school that did not make AYP on the basis on one subgroup (CEP, 2004). Concerns have also been raised about the validity of tests for some students with disabilities and ELLs and the appropriateness of holding all students to the same proficiency standards.

Although IASA and the 1997 amendments of the Individuals with Disabilities Education Act called for the inclusion of all students in assessment and reporting systems, states differed in whom they tested and how, and whether results for these students were reported and included in accountability calculations. Though states began to include more students with disabilities in state assessments through expanded accommodations and alternate assessments, few states included the scores of all tested students in their accountability calculations prior to the enactment of NCLB (Bolt, Krentz, & Thurlow, 2002; Thurlow, Lazarus, Thompson, & Robey, 2002). Regarding students with limited proficiency in English, some states excluded students who had resided in the United States or in their state up to 3 years and were enrolled in a bilingual or English as a Second Language (ESL) program. Other states exempted students based on the length of time spent in an ESL or bilingual education program or on their level of English proficiency (Goertz & Duffy, 2001).

The NCLB Act, however, requires states to test all students, regardless of their disability or proficiency in English, with reasonable adaptations and accommodations. States may not, as in the past, exempt ELLs based on

<sup>1</sup>The 14 states were Alaska, Arizona, California, Connecticut, Idaho, Louisiana, Maine, Montana, Nebraska, Nevada, New Hampshire, Pennsylvania, Utah, and Washington.

the length of their residency in the United States or on their level of English proficiency. These students may be assessed in their native language, although they must be assessed in reading/language arts in English after 3 years of attending a school in the United States. Most students with disabilities and all ELLs must be held to the same standards and proficiency targets as other students.

These provisions of NCLB raise several issues for states. The first issue concerns how students are tested and against what standards. For students with disabilities, this involves the use of alternative assessments and out-of-level testing—tests geared for grades lower than the grade attended by the student. The U.S. Department of Education initially did not allow states to set alternative standards for any students with disabilities, prohibited the use of out-of-level assessments, and limited the use of alternative assessments to 1% of all students (or about 8% of all students with disabilities). The department issued revised regulations in December 2003 to address, in their words, “misunderstandings about alternate assessments, alternate achievement standards, and the intent and purpose of the ... regulations” (Federal Register, December 9, 2003). The new regulations remove the cap on the use of alternative assessments. They allow states to establish alternative achievement standards for students with severe cognitive disabilities and to give some other students with disabilities out-of-level assessments. However, to discourage states from holding students to lower standards, only 1% of students in the grade levels tested can take tests based on alternative achievement standards or out-of-level tests *and* have their scores counted as “proficient” or “advanced” in calculating AYP, regardless of how many of these students meet proficiency levels. States may appeal the 1% cap, but only if their student bodies have larger populations of students with the most significant cognitive disabilities.

The assessment issue is different for ELLs who must be tested in reading (after their 1st year in the country) and mathematics, regardless of their English proficiency. Only about one third of the states offer versions of some of their assessments in languages other than English, generally in Spanish. Thus, many ELLs must take state assessments in their non-native language. Yet, research suggests that low ELL language ability decreases ELL performance on most tests, raising the issue of how valid these tests are in measuring content knowledge for these students. For ELL students, assessments in English become a measure of both subject matter and language. Some accommodations, such as modifying the linguistic complexity of test items or providing extra time, have the potential to reduce language barriers, but more research is needed on the validity of these changes (Abedi & Dietel, 2004).

A second issue involves how states treat special education students who take tests with nonstandard accommodations, those that result in a noncom-

parable score for the construct being assessed. These accommodations, which may be specified by a student's individualized education plan, can include having tests read to students, allowing students to use a calculator, having the test recorded for the student, and granting extended time for a timed test. States are struggling with how to include scores for these students in accountability calculations as they may not really reflect progress on standards-based learning constructs (Thurlow, 2003). Maryland, for example, invalidated the scores of students with disabilities who had test items on the reading portion of state assessments read to them. Students were given the lowest possible score for calculating AYP because items covering the decoding of text are not valid when they are read to students.

A third issue is the composition of the limited English proficiency (LEP) subgroup. Under NCLB, schools are expected to meet annual proficiency goals for LEP students. Educators, however, argued that this is an unrealistic requirement. Students exit the LEP subgroup once they achieve English proficiency and are replaced by students with no, or a limited, knowledge of English, making continuous improvement unlikely. After 2 years of discussing this problem, the U.S. Department of Education agreed to broaden the definition of the LEP subgroup beyond those students receiving direct services. States now have the option to include students who have attained proficiency in English for up to 2 years. Time will tell whether this adjustment addresses educators' concerns.

Finally, though states support the concept of disaggregating data to focus attention on the performance of student subgroups, they see less benefit in subgroup accountability mechanisms. Only half of the states report that setting annual performance targets for students with disabilities and ELL students will increase the academic performance of these students. Similarly, only half of the states feel that NCLB will narrow the achievement gap between special needs and nonspecial needs students. School districts also view narrowing these gaps as a major challenge (CEP, 2004).

### Capacity

The design of NCLB is based on the assumption that public reporting of test scores, the identification of schools that do not make AYP, and the threat of consequences for schools that fail to improve will create incentives for educators to work "harder and smarter." It also assumes that states and local school districts possess, or can develop, the capacity to assist school improvement efforts, to bring all students to proficiency, and to pay for these efforts. Mathis (2005/*this issue*) discusses two kinds of costs associated with NCLB: administrative costs of implementing the law, and

the costs of teaching children to standards. This section looks at another aspect of NCLB implementation—the existing capacity of states and districts to support school improvement—and at the fiscal capacity of states to fund school reform.

### *State and District Support for Low-Performing Schools*

Under NCLB, schools that do not make AYP for 2 years are identified for improvement. NCLB spells out a set of actions that states and districts must take with Title I schools identified for improvement and may take with non-Title I schools that are similarly designated. States and local districts are to share responsibility for supporting low-performing schools. States must establish statewide support systems composed of school support teams and of distinguished educators and principals from successful schools, a requirement of IASA as well. Local districts must ensure that low-performing Title I schools receive technical assistance in the development and implementation of their improvement plans, and, if necessary, implement corrective action for schools that fail to improve. Though IASA also required the preparation of school improvement plans, NCLB specifies that districts provide technical assistance in analyzing data, identifying and implementing effective professional development and instructional strategies, and analyzing and revising school budgets. Districts could design their own corrective action for schools that stayed in improvement status under IASA. NCLB defines what these actions should be, such as instituting a new curriculum with appropriate professional development, decreasing management authority or restructuring the internal organization of the school, appointing outside experts to assist the school, and extending the length of the school day or year.

Two national studies of NCLB—one by the CEP (2004) and a congressionally mandated study of Title I Accountability Systems and School Improvement Efforts (Shields, Lash, Padilla, Woodworth, & Laguarda, 2004)—show that districts, not states, have been the primary source of most kinds of assistance to low-performing schools. District assistance has focused primarily on school planning, disaggregating data, and curriculum adoption and alignment, and has been provided to most low-performing schools, regardless of Title I status. Schools that had been identified for improvement were much less likely to receive *resource-intensive* assistance through school support teams, full-time staff assigned to support teachers, and/or mentoring for the principal.

In 2002–03, 84% of districts with schools identified for improvement reported school planning or the use of student achievement data to plan improvement or monitor student progress as among their two most impor-

tant improvement strategies. Nearly all these districts assigned staff to help schools write their school improvement plans, analyze data to identify academic problems, choose research-based improvement strategies, and review budget and staff allocations. Though only one third of the districts identified the adoption of new curricula and instructional programs as a major change strategy, a majority of districts had required some or all of their schools to adopt new reading and language arts and/or mathematics curricula in the past 3 years. Often these new curricula were part of the districts' regular textbook adoption or curriculum revision cycle, but most districts viewed new adoptions as a strategy for improving and aligning instruction with state or district standards. A majority of districts took additional steps to align instruction by developing more specific local standards, publishing curriculum guides with standards, frameworks and pacing sequences, and issuing documents that mapped the content of required textbooks to standards and assessments (Shields et al., 2003).

About 60% of the districts with low-performing schools reported a major emphasis on increasing the quality and quantity of teacher and principal professional development, but all districts supported professional development in their schools, particularly in reading and mathematics instruction, curriculum and instructional alignment, and analyzing student data. Districts supported an array of professional development approaches, including teacher work groups, conferences, and in-class coaching and mentoring. Not all districts provided *intensive* support to identified schools, however. Only half of the districts provided school support teams and one third provided additional full-time school-level staff to support teacher development and/or mentors or coaches for the principal. Furthermore, this support was not available to all schools in improvement status. For example, in districts that provided these services, only 35% of identified schools reported they received help from their district's school support teams and only 39% received a principal mentor or coach. In some cases, districts targeted limited resources to their lowest performing schools; in others, low-performing schools volunteered to participate in district initiatives (Shields et al., 2004).

These averages mask considerable variation in the capacity of districts to assist their low-performing schools. In 2002–03, larger districts were more likely than smaller districts to report they provided resource-intensive support, such as school support teams, full-time staff developers, mentors for principals, and support for curriculum alignment. For example, about two thirds of the largest districts (more than 38,000 students) with identified schools said they provided full-time school-based staff to support teacher development compared to only 20% of the smallest districts (fewer than 3,500 students). Only half of small and rural school dis-

tricts provided any form of on-site assistance. Yet, the smallest districts had as many Title I schools identified for improvement as the largest districts (Shields et al., 2004).

This disparity in capacity is worrisome because most technical assistance comes from school districts. Districts report that they turn to multiple organizations for help: their state departments of education (94%), education service agencies or local consortia (70%), regional educational laboratories (49%), and institutions of higher education (47%) (CEP, 2004). However, as with districts, state assistance covers only a fraction of low-performing schools. In 2002–03, only one third of districts with identified schools reported that their state was the source of additional professional development staff or school support teams. Few districts received full-time staff to support teacher development or mentors for principals from their state (Shields et al., 2004).

Although states were required under IASA to develop school support systems for Title I schools in improvement, only 60% of the states reported in Summer 2003 that they included either school support teams or distinguished educators in these structures. These states provided different types and levels of support, however. Some states, like North Carolina, assigned a school support team to each school identified for improvement under the state accountability system and Title I. The teams, which underwent extensive training in data analysis, team building, resource alignment, and research-based practices, spent a year in most of these schools, assisting staff with the development and implementation of school improvement plans. Subject matter consultants also provided support to these and other schools. In South Carolina, teacher, curriculum, and principal specialists worked on site with low-performing schools for the entire academic year. Support teams and specialists in other states spent less time in identified schools, often only a few weeks in total, and focused their attention on conducting needs assessments and preparing school improvement plans.

Most states provided other types of support, such as technical assistance, professional development, assistance from regional support centers, and additional funding. Most states ran workshops for all Title I schools and for low-performing schools. Some states had full-time coordinators funded by Title I who assisted schools with needs assessments and improvement planning, and then connected schools to outside resources in their areas of need. Several states relied on their regional educational units to provide school improvement services that were funded through a combination of federal, state, and local dollars. Some states have reduced or eliminated their funding for these entities, however, leaving them increasingly dependent on contributions from local school districts. States also played a major role in pro-

viding school improvement grants to low-performing schools. About half of the districts reported that their identified schools received special grants from their states, most likely from state Title I school improvement funds (CEP, 2004). States often required that recipient schools use these funds for school improvement planning, reform initiatives, or professional development. Schools also relied heavily on Reading First and Comprehensive School Reform grants to support reform programs but again demand greatly outstripped the supply of these funds.

### *The Fiscal Context for Reform*

Helping schools and students meet the goals of NCLB is putting considerable strain on the human capacity of states and districts. Because of limited state and local resources, a minority of schools identified for improvement in 2002–03 received the kinds of intensive help needed to support instructional reform. At the same time, the new accountability provisions of NCLB have the potential to identify many more schools for assistance. Though only 6% of all schools were in program improvement status in 2002–03, an additional 22% of schools (approximately 20,000 schools) failed to make their 1st year of AYP (CEP, 2004). In addition, the cost of teaching all students to standards could require states to increase spending in the range of 20% to 40% (Mathis, 2005/this issue). This is all occurring at a time when the financial resources available to fund needed programs are sorely lacking.

In November 2002, the National Governors Association (NGA) reported that “states are facing a perfect storm: deteriorating tax bases, an explosion in health care costs, and a virtual collapse of capital gains and corporate profit tax revenues. The current problem is long-run and structural, and will take at least 3 to 5 years to remedy” (National Governors Association/National Association of State Budget Officers [NGA/NASBO], 2002). In the last 2 years, states have balanced their budgets through a combination of budget cuts, tax increases, and “one-shot” actions, such as spending surplus funds and borrowing against tobacco settlements and pension funds. State spending grew less than 1%, and state employment fell nearly 4% (NGA/NASBO, 2003). In the best cases, states provided small increases in, or froze, state aid to education. In the worst cases, states cut education aid. Overall, real per-pupil aid to school districts declined in 34 states between 2002 and 2004; in 19 states the decline exceeded 5% (McNichol & Harris, 2004).

As a result, some states cut back or scaled down their school assistance programs (Richard, 2004). Half of the states reported in 2003 that they did not have sufficient in-house expertise to provide technical assistance to all

of the schools and districts that had been identified for improvement, and another nine states felt they would be pressed to do so in the future (CEP, 2004). When faced with diminishing state aid, school districts downsized their own central office staff and cut programs and/or raised local taxes to maintain services and address growing enrollments.

Although the economy has begun to show signs of improvement, state finances remain fragile. State revenues, adjusted for inflation and legislated tax increases, have grown slowly since mid-2003. Personal income tax revenues are likely to grow far more slowly than they did in the late 1990s when they were buoyed by capital gains from the booming stock market. State surpluses are at an all-time low, and borrowing has created recurring obligations for years to come. The Congressional Budget Office projects that Medicaid spending will grow nearly 9% annually for the rest of the decade, faster than state revenue growth (NGA/NASBO, 2003). Legislatures are jittery about raising state taxes even higher. Yet, the Rockefeller Institute of Government cautions that, without further tax increases, it could take several years before state tax revenues reach the real per-capita levels of 2000 (Jenny, 2003).

The federal government has provided some fiscal relief, but it is short term. Congress enacted a package of \$20 billion in temporary relief for the states to be spread across 2003 and 2004. Funding for the Title I program jumped \$3.6 billion, or 41% between FY2001 and FY2004; but growth slowed to 3%, or \$400 million in FY2005; and President Bush's fiscal 2006 budget calls for only a \$600 million, or 5%, increase. Appropriations have fallen well behind Congressional authorizations for Title I, with a shortfall of more than \$13 billion in FY2002 through FY2004. In addition, 10 states and nearly one half of the nation's school districts will lose Title I funds in 2004–05 because their share of low-income children has declined (Title I Monitor, 2004).

## Conclusion

NCLB will have major consequences for how states and school districts hold schools accountable for student performance. The stringent requirements for participation in state assessments and the public reporting of disaggregated data and school performance will make disparities in student achievement much more visible. The question then becomes what actions states and school districts will and can take to address these problems given limited resources. It is likely that states will use their own school performance categories to continue to ration scarce school improvement resources. The lowest performing schools—under both state and federal def-

initions—will receive the most help, but research shows that it may not be of sufficient intensity to bring about the needed change. Until state revenues get back on track, states and local schools districts will be increasingly dependent on federal education aid to fund improvement activities. Yet, large increases in federal funding are unlikely in times of federal deficits.

State policymakers would like to see changes in NCLB that would provide greater flexibility in how they measure changes in performance (e.g., inclusion of growth models), include special needs students in accountability (e.g., increase the minimum *N* for special education and LEP students), apply sanctions (e.g., target sanctions to subgroups that fail to make AYP), and define highly qualified teachers. Congress will not consider amendments to the law until after the November 2004 elections, and major changes are unlikely. Although there has been a backlash against certain elements of the law and concern about insufficient federal funding, the bipartisan coalition that supported passage of the NCLB appears to be holding together. The civil rights community worries that amendments will weaken accountability for students of color and student with special needs. It remains to be seen, however, whether and to what extent the U.S. Department of Education will use regulations, guidance, and monitoring to hold the line or to offer relief to states and localities as the number of identified schools grows.

## References

- Abedi, J., & Dietel, R. (2004). Challenges in the No Child Left Behind Act for English language learners. *CRESST Policy Brief 7*. Retrieved April 20, 2004, from <http://cresst96.cse.ucla.edu/products/newsletters/policybrief7.pdf>
- Bolt, S., Krentz, J., & Thurlow, M. (2002). *Are we there yet? Accountability for the performance of students with disabilities* (Tech. Rep. No. 33). Minneapolis: University of Minnesota, National Center on Educational Outcomes. Retrieved April 15, 2003, from <http://education.umn.edu/NCEO/OnlinePubs/Technical33.htm>
- Center on Education Policy. (2004). *From the Capitol to the classroom: Year 2 of the No Child Left Behind Act*. Washington, DC: Author.
- Driscoll, W., & Fleeter, H. (2003). *Projected costs of implementing the federal "No Child Left Behind Act" in Ohio* (prepared for the Ohio Department of Education). Columbus, OH: Author.
- Federal Register. (2003, December 9). Title I: 34 CFR Part 200: Sec. 200.13 as adopted in the final regulations published in the Federal Register (68 FR 68697). Retrieved April 2, 2004, from <http://wais.access.gpo.gov>
- Goertz, M. E. (2001). Standards-based accountability: Horse trade or horse whip? In S. H. Fuhrman (Ed.), *From capitol to the classroom: Standards-based reform in the states* (pp. 39–59). One hundredth yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press.

- Goertz, M. E., & Duffy, M. C. (2001). *Assessment and accountability systems in the 50 states: 1999–2000*. RR–046. Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.
- Hoff, D. J. (2004, March 10). Accountability conflicts vex schools. *Education Week*, 23(26), 1. Improving America's Schools Act of 1994, Pub. L. No. 103–82.
- Jenny, N. W. (2003, November). State tax revenue shows slight improvement. *The Rockefeller Institute State Fiscal News*, 3(8). Retrieved December 3, 2003, from <http://StateAndLocalGateway.rockinst.org>
- Kossan, P. (2003, September 29). 220 schools in Arizona on feds' underachieve list. *The Arizona Republic*. Retrieved July 23, 2004, from <http://www.azcentral.com/arizonarepublic>
- Mathis, W. (2005/this issue). The cost of implementing the Federal No Child Left Behind Act: Different assumptions, different answers. *Peabody Journal of Education*, 80(2), 91–120.
- McNichol, E., & Harris, M. (2004). Many states cut budgets as fiscal squeeze continues. *Center on Budget and Policy Priorities*. Retrieved April 23, 2004, from [www.cbpp.org/4-22-04sfp.htm](http://www.cbpp.org/4-22-04sfp.htm)
- National Assessment of Title I, Independent Review Panel (NATI/IRP). (2001). *Improving the odds*. Washington, DC: Council for Basic Education.
- National Governors Association/National Association of State Budget Officers. (2002, November). *The fiscal survey of states*. Retrieved April 9, 2003, from [www.nasbo.org/publications/fiscalsurvey/nov2002fiscalsurvey-revisedc.pdf](http://www.nasbo.org/publications/fiscalsurvey/nov2002fiscalsurvey-revisedc.pdf)
- National Governors Association/National Association of State Budget Officers. (2003, December). *The fiscal survey of states*. Retrieved December 28, 2003, from [www.nasbo.org/publications/fiscsurv/fsfall2003.pdf](http://www.nasbo.org/publications/fiscsurv/fsfall2003.pdf)
- No Child Left Behind Act of 2001, Pub. L. No. 107–110.
- Olson L. (2003, December 10). In ESEA wake, school data flowing forth. *Education Week*, 23(15), 1.
- Richard, A. (2004, January 7). States unable to help all struggling schools. *Education Week*, 23(16), 19.
- Schema, D. J. (2004, March 25). 14 states ask U.S. to revise some education law rules. *New York Times*. Retrieved March 30, 2004, from <http://www.nytimes.com/2004/03/25/education/25CHIL.html>
- Shields, P. M., Laguarda, K. G., Lash, A., Lopez, A., Padilla, C., Skolnik, H., et al. (2003). *Evaluation of Title I accountability systems and school improvement efforts (TASSIE): Second-year Technical Report* (draft). Palo Alto, CA: SRI International.
- Shields, P. M., Lash, A., Padilla, C., Woodworth, K., & Laguarda, K. G. (2004). *Evaluation of Title I accountability systems and school improvement efforts (TASSIE): Second-year findings*. Washington, DC: U.S. Department of Education.
- Thurlow, M. L. (2003). Biting the bullet: Including special needs students in accountability systems. In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 115–137). New York: Teachers College Press.
- Thurlow, M. L., Lazarus, S., Thompson, S., & Robey, J. (2002). *2001 state policies on assessment participation and accommodations*. (Synthesis Rep. No. 46). Minneapolis: University of Minnesota, National Center on Educational Outcomes.
- Title I Monitor. (2004, July). Census updates slash Title I funding for over half of nation's school districts. *Title I Monitor*, 9(7), 1, 8–15.
- U.S. Department of Education, Planning and Evaluation Service. (2001). *High standards for all students: A report from the National Assessment of Title I on progress and challenges since the 1994 reauthorization*. Washington, DC: Author.
- U.S. General Accounting Office. (2003). *Title I: Characteristics of tests will influence expenses; information sharing may help states realize efficiencies*. GAO-03-389. Washington, DC: Author.