

Related Brookings Resources

- *The Brown Center Report on American Education*, Tom Loveless, editor (published annually)
- *Brookings Papers on Education Policy*, Diane Ravitch, editor (published annually)
- *Who's in Charge Here? The Tangled Web of School Governance and Policy*, Noel Epstein, editor (Brookings Institution Press and the Education Commission of the States, 2004)
- *No Child Left Behind? The Politics and Practice of School Accountability*, Paul E. Peterson and Martin R. West, editors (2003)

To receive a weekly e-mail about Brookings news, events, and publications, sign up for the Brookings Alert at www.brookings.edu.



1775 Massachusetts Ave., N.W.
Washington, DC 20036

No Child Left Behind: How to Give It a Passing Grade

MARTIN R. WEST

The No Child Left Behind Act (NCLB) has the potential to improve many of America's schools, but this potential is currently undermined by serious flaws in how the

program evaluates school performance. Because NCLB's measurement system compares only students' performance at a single point in time against state-determined standards, the information generated on school performance is often misleading and creates perverse incentives for states to lower their expectations.

Fortunately, new measures of school performance based on the academic progress made by individual students over time offer a promising alternative to the law's current approach. And a recently announced NCLB pilot program allowing up to ten states to use growth-based school accountability models represents an important step toward making the law more effective. But the pilot's effectiveness will depend ultimately upon whether the Department of Education allows states sufficient flexibility in devising such alternative accountability schemes.



AN NCLB PRIMER

NCLB's overarching goal is that all American students reach math and reading proficiency by 2014. As a condition of receiving federal aid, the law places three key requirements on the states: to assess the performance of all students annually in grades three through eight and once in high school against state-determined proficiency standards in math and reading; to disclose the results to the public; and to sanction and eventually intervene in schools and

districts where students (or one of several student subgroups) fail to meet statewide performance goals. In short, NCLB mandates that states adopt comprehensive accountability systems for identifying and improving underperforming schools.

Less than four years after the path-breaking law's passage, it remains too soon to assess definitively its impact on student achievement. There is little doubt, however, that by providing a wealth of new



information about the performance of students against state standards, NCLB has shined a light on ethnic and racial disparities in achievement in both urban and suburban schools, while creating new pressure for reform and innovation. The law has thus maintained the support of the Education Trust, the Council of Great City Schools, the Citizens' Commission on Civil Rights, and other respected advocates for disadvantaged students from across the political spectrum.

Yet NCLB also has its detractors. The teachers unions and a growing number of states and school districts are working actively to soften key provisions, to use them as leverage to extract more federal dollars for public education, or to do both at once. Congressional Democrats allege that the Bush Administration has not spent the money it promised on the law's implementation, even as legislatures in several Republican-dominated states complain of an unwarranted federal intrusion.

While some of these objections are no doubt questionable, another line of criticism must be taken more seriously. Many policy-makers and researchers who are sympathetic with the law's goals and its emphasis on annual testing are nonetheless unhappy with NCLB's specific approach. Officials in various states contend that if allowed greater flexibility, they could do a better job of determining which of their schools are improving student achievement and narrowing gaps between low-performing and high-performing student subgroups. Unfortunately, their particular concerns are frequently ignored in mainstream discussions.



Martin R. West is a research fellow in the Governance Studies program at the Brookings Institution.

AN UNFUNDED MANDATE?

Indeed, public debate over NCLB has focused not on the substance of the law's accountability system, but rather on the money allocated to put it in place. A point of contention in the 2004 presidential campaign, the issue has remained in the public eye in part as a result of separate lawsuits filed in 2005 by the National Education Association and the state of Connecticut, lawsuits alleging that the law is an unfunded mandate and requesting relief. Are such allegations credible? More to the point, is more federal money what is needed to improve American education?

Set aside the fact that NCLB is not, by any legal definition, an unfunded mandate. Because the money is offered to states as a grant-in-aid, states are free to turn it down if they dislike the strings attached. Claims that the law is underfunded also fail on the merits. As the nonpartisan Government Accountability Office has shown, test-based accountability is an intrinsically inexpensive reform strategy. Nationwide, cost estimates run as low as \$9 per student, on average, for the type of tests currently used, and nearly all independent estimates of the costs of testing come to less than \$50 per student out of the roughly \$10,000 per student currently being spent on their education. Nor have the law's provisions requiring that students in persistently failing schools be offered public school choice and supplemental services yet placed much of a fiscal burden on states and school districts.

But to delve too deeply into the debate over funding levels is to miss the forest for the trees. Simply put, America's educational woes have little to do with the amount we

spend on the public schools. International comparisons place us at the very top in per pupil expenditure but near the bottom (24th out of the 29 OECD countries participating in a 2003 assessment of mathematics literacy among 15-year-olds, to cite one recent example) in terms of student achievement. Moreover, while school spending, adjusted for inflation, has more than doubled since 1970, student achievement remains disappointing, and high-school dropout rates have increased. Of course, this pattern is not determinative: More money may ultimately be needed to achieve NCLB's lofty goals for student performance. Because no national school system has achieved near-universal proficiency in core academic subjects, we simply don't know what doing so will require. Yet it is foolish to invest evermore resources into a failing system that has shown few signs of improvement.

NCLB was drafted with this record in mind. By requiring states to establish a rigorous accountability system as a condition for receiving federal funds, it aims to convert federal aid for education from a subsidy for state school systems as they currently exist into a lever for making those systems more effective and more equitable. The generally positive results seen from test-based accountability systems – states that adopted such systems in the 1990s significantly improved their relative standing on the federally administered National Assessment of Educational Progress – suggest that the law's general approach makes sense.

A FLAWED MEASURING STICK

But the effectiveness of accountability systems in education, as in other fields,

depends in the first instance upon the accuracy of their performance measures. If a school accountability system identifies schools where student learning flourishes, it can provide useful information to parents, teachers, administrators, school board members, state policy-makers, and the public at large. The most pressing shortcomings of the NCLB accountability system therefore involve the measuring methodology states are required to use. NCLB's current method for assessing school quality provides misleading information about schools both within and between states.

Within States. Under NCLB as it is currently implemented, states must evaluate schools based primarily on their students' performance at a single point in time. Schools are said to be making "adequate yearly progress" when their students (and all student subgroups above a minimum size) meet statewide targets for the percentage of students who are proficient according to state-determined standards. States must raise these targets at regular intervals until 2014, by which time all students are expected to be proficient.

Unfortunately, this level-based accountability system provides little information about how much students in a school are actually learning each year. In fact, according to my research on Florida's school system, the gains in reading performance made by students attending schools that made adequate yearly progress during the 2003-04 school year were, on average, no larger than the gains made by students in schools not making adequate yearly progress. In math, students in schools making adequate yearly progress made gains that were 3 percent of a standard deviation larger, a negli-

"The most pressing shortcomings of the NCLB accountability system involve the measuring methodology states are required to use."



“Indeed, it is the states with the highest expectations for their students that are most likely to be found lacking under the federal law.”

gible difference. At least in Florida, which is unique in the quality of data it has made available to investigate these issues, a school's rating under NCLB seems to have more to do with the composition of its student body than the progress its students were making in the classroom. Schools not making adequate yearly progress in the law's second year had, on average, 40 percent more poor students and a substantially greater share of minorities.

The problem is that the NCLB methodology for measuring school performance does not pay enough attention to the vast differences in students' academic preparation when they arrive at school – differences that have clear consequences for their subsequent test scores. Schools with large numbers of disadvantaged students can be deemed failing for not meeting statewide proficiency targets even if their students are making dramatic progress. Conversely, schools in affluent communities may appear to be effective despite the fact that their students are learning less than the state's average student from one year to the next.

Seeking to compensate for these problems, a few states and school districts have developed alternative measures of performance, termed growth models, that incorporate information on where students began the year in addition to where they end up. Not surprisingly, these measures often provide a quite different picture of schools' performance. In Florida, for example, 62 percent of the state's schools did not make adequate yearly progress in the 2004-05 school year. But more than a third of those failing schools did well enough to earn an A or B on the state's 5-category school grading system, which awards half of its points based on the percentage of students who improved their performance against state standards over the

previous year. Local officials in Florida, who have no stake in either the state or federal accountability system, almost uniformly contend that the state's growth model does a better job of identifying effective schools.

Between States. But the difficulties of the NCLB measuring methodology extend beyond the borders of particular districts and states. Because NCLB allows states to create their own tests and to define the level of achievement required for students to be deemed proficient, states vary widely in their expectations of what students should know. The share of students in a state who are proficient therefore contains little information about the relative effectiveness of its schools. Indeed, it is the states with the highest expectations for their students – most of whom set their standards before NCLB's passage – that are most likely to be found lacking under the federal law.

Results from the 2003 National Assessment of Educational Progress (NAEP), a previously voluntary national exam that NCLB requires states to administer, demonstrate how widely expectations vary from state to state. Incredibly, the results show that there is almost no relationship between the percentage of students in a state deemed proficient according to state standards and the percentage reaching proficiency on the NAEP. Students in Texas and South Carolina, for example, performed similarly against national norms, with just over one quarter of students reaching proficiency in reading. However, fully 83 percent of students in Texas achieved proficiency in 2003 on the state's own exam, as compared with 29 percent of students in South Carolina.

As many observers have noted, the ability of states to alter their standards raises the specter of a nationwide race-to-the-bottom, with states progressively lowering their expectations for students so that fewer schools are identified as failing. Indeed, a handful of states including Louisiana, Colorado, Connecticut, and Arizona have altered their scoring systems since the law's passage in an apparent effort to increase the number of schools making adequate yearly progress.

Why did NCLB's Congressional authors settle for an accountability system with such seemingly predictable deficiencies? Unworkable compromises often emerge from legislative hoppers, and, in this case, any move toward establishing national standards had to travel a particularly rocky road. Prior efforts to create national standards had floundered on attempts to define them. As former Assistant Secretary of Education Chester Finn quipped in the wake of a failed 1990s attempt to accomplish the task, "Republicans oppose any proposal with the word 'national' in it, Democrats oppose anything with the word 'standards.'" In the case of NCLB, establishing a full-fledged accountability system for schools at the same time only augmented the problem. Meanwhile, the idea that all schools, no matter what the composition of their student body, should be held to a common standard, rather than be evaluated against their own performance the year before, resonated with the law's rhetorical commitment to the notion that all students can learn.

What is less widely recognized is that by 2001 only a handful of states even had the capacity to measure the annual progress of

individual students—the most basic requirement for a school accountability system based on the growth in their achievement. Although many states have since upgraded their data systems, a recent survey by the non-profit Data Quality Campaign revealed that most still lack the necessary resources to move immediately to a growth-based model.

IMPENDING COLLAPSE

The problems with the NCLB measurement system are rapidly coming to a head. Consider the following scenario: Higher statewide performance targets cause a sharp increase in the percentage of schools failing to make adequate yearly progress, just as state accountability systems register conflicting signals about school performance, some indicating considerable progress. An *Education Week* analysis of preliminary data from the 2004-05 academic year showed just such an outcome unfolding in several states. In California and Hawaii, for example, the percentage of schools making adequate yearly progress decreased by 10 and 21 percentage points, respectively, over the previous year, despite the fact that the percentage of proficient students in each state increased. In Hawaii, the percentage of schools making adequate yearly progress fell to 34 percent, the lowest of any state to have reported its data thus far. But the local reaction was not entirely negative. "We have two daily papers," Hawaii's communications director explained. "One played it up like the glass was half-empty; the other like the glass was half-full. So it's kind of confusing." Such schizophrenic outcomes could lead many people to question the legitimacy of the entire accountability enterprise. After all, if virtually all schools in a state are identified as failing—including many that appear to be



“Growth-based accountability systems have the potential to offer a fairer and more accurate assessment of school performance.”

succeeding in difficult circumstances—is the problem with the schools or with the accountability system? The threat of diminished credibility is especially acute in places like Florida where, as we have seen, dual school-rating systems provide conflicting assessments of the effectiveness of specific schools.

The Department of Education is hardly unaware of these dangers. Its strategy to date, introduced by Secretary of Education Margaret Spellings in April of 2005 as a “new, common-sense approach to implementation,” has consisted mainly of allowing states to make minor modifications to their accountability plans, apparently in the hopes of postponing the day of reckoning until school performance improves. Various states have been allowed, for example, to delay scheduled increases in their performance targets, to use a larger minimum number when determining whether the performance of a subgroup of students within a school will be assessed separately, or to make statistical allowances for the uncertainty inherent in any measure of school performance.

Individually, each of these changes has been reasonable and even prudent, given the circumstances. And they seem to have helped prevent a dramatic increase in the share of schools not making adequate yearly progress in the 2004-05 school year. But their collective effect has been the creation of a patchwork system in which the apparent success of a state’s schools under NCLB depends as much on the savvy and sophistication of the statisticians in its education agency as it does on the performance of its schools. Differences in the federal treatment of states requesting flexibility even provided ammunition for Connecticut,

whose allegations against Secretary Spellings in court include the claim that her department’s enforcement of NCLB has been arbitrary. Meanwhile, the various modifications have done nothing to ensure that the schools identified as making adequate yearly progress are those in which students are actually learning.

THE GROWTH-MODEL PILOT PROGRAM

Against this backdrop, the Department of Education’s November 18, 2005 announcement of a new growth-model pilot program represents an important step forward; indeed, it is the most important regulatory change since the passage of NCLB. The program will allow up to 10 states to implement accountability systems based in part on annual “growth” in student achievement – that is, the amount individual students are learning from one year to the next, as measured by the state achievement test. The first growth models may be approved for use in the 2005-06 academic year, well in advance of the law’s scheduled reauthorization in 2007.

Such growth-based accountability systems have the potential to offer a fairer and more accurate assessment of school performance. Their widespread adoption could also reduce pressures to lower state proficiency standards, preventing the potential dumbing-down of school curricula by rewarding schools for gains made by high-achieving students. In states with the necessary database capacities to participate, the pilot program should help sustain support for the law among officials frustrated with the federal government’s hitherto rigid approach to implementation. And, by encouraging other states to invest in

such data systems now, it may help ensure that Congress is less constrained than it was in 2001 by what states can do when the law comes up for reauthorization.

It remains unclear, however, whether the Department of Education will allow states enough flexibility to make the pilot as informative as it should be. While there is an emerging consensus among practitioners and scholars that measures of school performance based on growth are superior to level-based measures, there is little agreement over how best to implement them in the context of an accountability system. Growth models bring with them a host of technical and political problems that lack broadly accepted solutions. To address these issues, researchers will need a solid base of evidence on how various growth-based accountability systems work in practice.

That said, the Department of Education has been wise to exclude one popular category of growth models – those commonly referred to within education circles as “value-added” models – from the pilot program. Value-added models incorporate information on students’ background characteristics when evaluating their progress and, as a consequence, have been appropriately criticized for reintroducing and disguising lower expectations for disadvantaged students. Secretary Spellings should also insist that states experimenting with growth-models continue to report test score levels to the public separately by subgroup as mandated under the current system.

Participating states should otherwise be given considerable flexibility, including

the flexibility to use growth-models that are not premised on the notion that every student, regardless of his or her grade, will be fully proficient by 2014. The federal government should instead allow states to reward schools for putting virtually all students on a trajectory that, if sustained, will ensure that they are fully proficient by the time they are tested in high school. While this new interpretation of the law’s language on deadlines would be characterized by some as a step back, it is increasingly clear that the requirement that all students in a school be proficient by 2014 will, sooner or later, undermine the credibility of the entire accountability system. Or it will lead to a state-by-state downward redefinition of the meaning of proficient.

CONCLUSION

When considering how much flexibility states should be allowed under NCLB, whether in its current form or after its reauthorization, it is useful to recall the role states played in the law’s initial development. The accountability movement in education was a state-led effort, with the crucial steps taken by governors eager to establish a reputation for reform. Likewise, the law’s core principles of annual assessment and disaggregation of achievement data by subgroup did not emerge in whole cloth from the federal legislative process leading up to NCLB. Rather, these principles were developed independently by a few innovative states, most notably Texas and North Carolina, and gained credibility when those same states’ performance on the National Assessment of Educational Progress improved dramatically. We should again let the states lead the way.

POLICY BRIEF



Recent Policy Briefs

- "Treatment of Coronary Artery Disease: What Does Rationing Do?"
Henry J. Aaron (December 2005)
- "Health Care Rationing: What it Means,"
Henry J. Aaron (December 2005)
- "Why Federalism Matters,"
Pietro S. Nivola (October 2005)
- "Moving Toward Smarter Aid,"
Lex Rieffel (August 2005)

Editor
Adam Wolfson

Production/Layout
Gladys L. Arrisueño
Merrick Buchanan
Bethany C. Hase

**The Brookings Office
of Communications**
202/797-6105
communications@brookings.edu

The views expressed in this Policy Brief are those of the author and are not necessarily those of the trustees, officers, or other staff members of the Brookings Institution.

Copyright © 2005
The Brookings Institution

There was always a danger in the highly prescriptive nature of NCLB regulations that whatever good was accomplished by bringing some states into the accountability fold would be more than offset by the prevention of experimentation and innovation. By granting flexibility only to states that have proven themselves to be

leaders in the effort to increase accountability in education, the growth-model pilot program provides a way to eliminate this tradeoff. The Department of Education should trust these states to serve as "laboratories of accountability" with the aim of devising new and better measurement systems. **B**

**Tell us what you think of this Policy Brief.
E-mail your comments to feedback@brookings.edu.**

The Brookings Institution
1775 Massachusetts Ave., NW
Washington, DC 20036

| |
|---|
| NONPROFIT ORG. U.S. POSTAGE PAID FREDERICK, MD PERMIT NO. 225 |
|---|