1

Introduction: Can Standards-Based Reform Help Reduce the Poverty Gap in Education?

ADAM GAMORAN

Pervasive inequality is the most pressing problem facing U.S. education. While average achievement levels in some U.S. school districts equal those in the world's high-achieving nations, other districts rank among the world's low performers. Inequality is evident not only between districts but also within districts and within schools, where students of different social backgrounds attain widely varying outcomes. The problem is particularly pronounced for students who face economic disadvantages. While students from disadvantaged racial and ethnic groups made noteworthy progress over the last forty years (mainly from around 1970 to 1990), gaps among students from families with varied economic resources remained stable and wide throughout that period. Achievement differences between students living in poverty and their more privileged peers, often called the "poverty gap," have shown little sign of diminishing.

The current federal program designed to reduce inequality in education, the No Child Left Behind Act of 2001 (NCLB), is the latest in more than two decades of federal efforts to raise educational standards. From *A Nation at Risk* in 1983 to the National Goals Education Panel of 1990 and the Goals 2000 Act in 1994, federal policy had attempted to increase standards and better align curriculum, instruction, and assessment in the nation's schools. NCLB is unique among the government's efforts in that it focuses not only

on raising standards overall but also on increasing the achievement of students in a variety of demographic subgroups, including those from racial and ethnic minorities, disabled students, English language learners, and students faced with economic disadvantages. NCLB also is much stronger and more far-reaching than previous federal efforts to raise education standards. For example, to receive federal education aid, states must set standards for student performance and every year assess reading and mathematics achievement for students in third through eighth grades and in high school. In selected grades, science achievement must also be assessed. In addition, at least 95 percent of students in each school and district must take the tests; no more than 3 percent of disabled students may take an alternative assessment; and English language learners must take the assessments within two years of arrival in the United States. Those measures prevent schools from "hiding" their low achievers and thus inflating their results. Failure to meet the achievement targets (known as "adequate yearly progress," or AYP) results in sanctions that range from requiring districts to offer students tutoring and transfer options to closure and reconstitution of schools.

There are many reasons to expect that NCLB's approach to increasing standards and holding schools accountable for student performance will boost the chances for poor children to succeed in school. First, by requiring schools to report test results separately for students in different demographic subgroups, NCLB shines a spotlight on social inequalities in school performance that sometimes have been obscured in the past, perhaps increasing the political will to address this profound problem. Second, in principle, the transfers and supplemental services offered to students in schools that are not making AYP should help disadvantaged students to obtain better opportunities. Third, NCLB requires districts to place a "highly qualified teacher" in every classroom. Teaching out of field or with provisional certification is more common in schools with large proportions of low-income students than elsewhere, so that requirement also may improve opportunities for the disadvantaged. Finally, NCLB requires districts participating in the Reading First program to choose curricula and teaching methods for which there is scientific evidence of success. If those methods are more effective than untested alternatives, the move to evidence-based teaching of reading may be especially important for poor children, who are overrepresented among struggling readers.

Yet there are many challenges to reducing inequality under NCLB. More diverse schools may be more likely to be labeled as not making adequate yearly progress simply because their larger number of population subgroups

means that they have more targets to hit. The massive level of improvement required by 2014 also suggests that schools with high concentrations of disadvantaged students may be unable to succeed. NCLB sanctions such as transfers and supplemental services draw resources *away* from struggling schools, so students who are not fortunate enough to transfer may find their opportunities getting worse, not better. Indeed, the students least likely to transfer when their schools are failing may be those living in poverty whose parents have low levels of education. Moreover, even if districts provide extra resources to schools with low-income students (for example, through the Reading First program), those resources may not be enough to compensate for a disadvantaged home and community environment. Also, many districts are passing accountability for meeting standards to students themselves, and some (but not all) research suggests that students living in poverty are disproportionately burdened with sanctions when accountability systems are put in place. Finally, NCLB confronts a variety of political challenges, including resistance to implementation and bias in implementing Reading First (as revealed by the U.S. Department of Education's Office of Inspector General in reports such as *The Reading First Program's Grant Application Process*, which appeared in September 2006). Resistance and bias in implementation may prevent NCLB from reducing inequalities even if the theories behind the legislation are valid.

Many of the questions raised in this volume are amenable to empirical analysis. While it is too soon to assess outcomes under NCLB, standards-based reform by states and school districts has been in place long enough to allow researchers to measure its impact, and their findings can be used to guide the implementation and possible revision of NCLB. To address the possibilities, the Institute for Research on Poverty and the Wisconsin Center for Education Research sponsored a conference at the University of Wisconsin–Madison in February 2006. Participants came from a variety of disciplinary backgrounds, including economics, political science, psychology, and sociology, and their presentations included syntheses of work to date and new empirical studies. Through careful analysis and lively debate, speakers, discussants, and audience members sifted through the evidence to assess the relation between NCLB and the poverty gap. Though complex and sometimes contradictory, the findings pointed toward modest improvements for poor children, but at nowhere near the rate of improvement demanded by NCLB—or by anyone who views low educational outcomes among disadvantaged youth as a major impediment to the advancement of American society.

Schools, Standards, and Gaps in Achievement

NCLB is no doubt flawed by its assumption that schools alone can eliminate achievement gaps in the face of powerful social inequalities in the wider society. Yet while that assumption is surely unrealistic, the question of what schools *can* do is unresolved. How much gap-closing can be expected from standards-based reform? What lessons can be learned from past reform efforts that will lead to greater progress, if not fully accomplish NCLB's ambitious goals? On one level, NCLB creates incentives for improving student performance and reducing gaps in achievement, and on that level, the policy appears to be succeeding. To an extent never before attained in the United States, educators, politicians, and the general public have been alerted to the problems of inequality, including inequality between students in poverty and their more advantaged counterparts. Incentives, however, are unlikely to suffice. Instead, specific strategies are needed to improve students' learning opportunities in schools, such as strategies to improve teacher and instructional quality and to promote evidence-based practices. Moreover, to make a difference in achievement gaps, those strategies need to target the most disadvantaged students. To what degree has standards-based reform led to improvements in the quality of teachers and teaching? If improvements are evident, have they occurred in schools with disadvantaged student populations? Those questions are addressed in chapters 2 through 4 in this volume.

The main targets of NCLB sanctions are school districts and schools. In response, many states and districts are targeting students, reasoning that if principals, teachers, and support staff are working hard to raise standards, students should respond with greater efforts of their own. Common sanctions include not promoting students who perform poorly on assessments to the next grade and not allowing them to graduate from high school if they fail the exit exam. While grade retention and graduation testing are not included in NCLB, they are increasingly being employed alongside NCLB. What can be learned from recent experiences with sanctions of students that will aid understanding of the prospects for reducing achievement gaps under NCLB? Chapters 5 and 6 address that question.

School choice and supplemental tutoring are among the key sanctions that NCLB places on schools that fail to bring students to the required level of achievement. While those sanctions clearly are galvanizing attention, the empirical question is whether they make a difference in student achievement. Given the experience before and during the early implementation of NCLB, should achievement gaps be expected to narrow as schools and districts

implement the sanctions, as required under NCLB? To answer that question, which is addressed in chapters 7 and 8, one must attend both to the impact of the strategies when they are implemented well and to the extent and quality of their implementation as NCLB moves forward.

To draw lessons for NCLB from recent experiences under standards-based reform, it is essential to understand the political context of NCLB and to examine the early implementation of the legislation. Those topics are addressed in chapters 9 and 10, which look to NCLB's future. NCLB is scheduled for reauthorization in 2007, but many observers predict that it will be delayed until after the 2008 presidential election. The findings in this book are intended to inform the debate over reauthorization, whether immediate or delayed.

Findings on Standards-Based Reform and Inequality

Following this introduction, chapter 2 provides a background for understanding the relation between standards-based reform and the poverty gap. Barbara R. Foorman, Sharon J. Kalinowski, and Waynel L. Sexton draw attention not only to NCLB and its immediate predecessors in federal education policy but also to the civil rights movement, the War on Poverty, and the broader recognition of the rights of all students to a quality education as a means of ameliorating social disadvantages. With her coauthors, Foorman, who brings special insight to research-based understanding of NCLB from her recent service as first commissioner of education research at the Institute of Education Sciences, shows that the same milieu that gave rise to standards-based reform also stimulated the right to education movement, which supports quality education for students with disabilities. Until recently, laws regarding general education improvement and laws regarding special education have been implemented separately. With the passage of the 2004 Individuals with Disabilities Education Act (IDEA), however, the possibility emerged of a coordinated approach to ensuring high-quality schooling for all children, regardless of their social background or disability. Foorman and her colleagues recognize the tensions between IDEA and NCLB, but they maintain that those tensions can be resolved.

While IDEA and NCLB offer important new opportunities to reduce the poverty gap, Foorman and her colleagues argue that they represent only the first step. The quality of the implementation of the measures is crucial to their success. In the view of the authors, key aspects of implementation include the use of instruction based on scientific evidence of effectiveness,

the alignment of instruction with standards and assessment, and the application of effective measures early in children's lives.

Findings on Accountability Reforms and the Quality of Teachers and Teaching

Among the major elements in NCLB's strategy for raising test scores and reducing test score gaps is improvement in the quality of teachers and of teaching. In chapter 3, Meredith Phillips and Jennifer Flashman set the tone for the empirical work in this book with their examination of NCLB-like policies during the 1990s. Recognizing that it is too soon to evaluate NCLB directly, they focus on how changes during the 1990s in state accountability policies were linked to changes in teacher and instructional quality within states. They consider not only potential positive consequences, such as those anticipated by the designers of NCLB, but also unintended, negative consequences, such as the possibility that the rigors of stricter accountability may drive the best-qualified teachers away from the profession, or, in the case of the 1990s, to states with weaker accountability regimes. They use accountability reports from the Council of Chief State School Officers to assess changes in state policies and the Schools and Staffing Surveys of 1993–94 and 1999–2000 to measure state changes in teacher and instructional quality. They estimate state fixed-effects models to rule out changes attributable to fixed, unobserved attributes of states.[1]

Overall, Phillips and Flashman find that variation in state accountability policies was not consistently linked to changes in teacher and instructional quality. Their evidence offers little basis for the concern that accountability policies may be detrimental to teacher quality. They observe that sanctions directed at teachers tend to reduce perceptions of autonomy among teachers in high-poverty schools but that that tendency might indicate that teachers are being pushed to align instruction with standards, as intended under standards-based reform. They also find, not surprisingly, that in states that increased the use of student testing for accountability, teachers participated more often in professional development devoted to student assessment; teachers also devoted more instruction time to subjects that were more heavily tested. Both findings illustrate the power of standards-based reforms to affect teachers' behavior. Phillips and Flashman further observe that accountability measures directed at teachers were associated with a variety of positive

1. State fixed-effects models eliminate all variation associated with fixed characteristics of states by including an indicator variable for each state or by subtracting the state mean value from each data point. Such models require data on states at more than one point in time.

changes, including increases in teacher experience, certification, college quality, and rates of teachers holding advanced degrees and remaining in the profession. However, those findings did not hold for all types of schools. Some were evident in schools attended mainly by disadvantaged students, while others appeared in schools with more advantaged populations; there was no clear pattern. Moreover, while the findings seem promising, it is striking that the most persistent effects were associated with policies directed at teachers, not schools, whereas NCLB mainly targets schools and reaches teachers only indirectly. Consequently, it will be important to monitor NCLB's impact to determine whether it needs to aim more directly at teachers to affect the quality of teachers and teaching.

Teacher and instructional quality also are central to chapter 4, by Laura M. Desimone, Thomas M. Smith, and David Frisvold. The authors draw on teacher data from the 2000 and 2003 National Assessments of Educational Progress (NAEP) to examine changes over a slightly later, albeit shorter, period of time. The NAEP teacher data, which are representative of participating states, include information not only on background and training but also on approaches to instruction, as reported by the teachers themselves. The authors consider whether teacher characteristics and practices differ in high-poverty and low-poverty schools and, within schools, for students who receive free lunches and those who do not. They also examine whether the differences observed changed from 2000 to 2003 and whether any changes were linked to state accountability policies, which they measured using a variety of data sources on the dimensions of power, authority, consistency, specificity, and stability of standards. After a series of descriptive findings, they also estimate models with state fixed effects.

Consistent with past research, the findings of Desimone and her colleagues show that in 2000, low-poverty schools tended to have teachers with better credentials and math backgrounds than high-poverty schools. Within-school gaps also were evident: disadvantaged students had less access to qualified teachers in both high- and low-poverty schools. The gaps changed little between 2000 and 2003. Moreover, most of the state policy variables were not linked to changes in teacher quality. One exception, however, was that states that increased their alignment of standards and assessments (consistency) witnessed small increases in the percentage of certified teachers and of teachers with mathematics majors. Finally, the authors report two notable changes in teacher quality related to school poverty levels. On the one hand, states that began ranking schools (an indicator of power) narrowed the gap between high- and low-poverty schools in the likelihood of having certified

teachers. On the other hand, states that introduced more specific standards saw *increases* in the gap between high- and low-poverty schools in teacher certification levels. Because of the short time frame, the findings are tentative, but they suggest that publicizing schools' success or failure in reaching standards may help bring schools attended by students in poverty closer to their more advantaged counterparts—as intended by the designers of NCLB—whereas excessive specification of standards may make it more difficult for high-poverty schools to recruit certified teachers.

Findings on Sanctions for Students: Grade Retention and Exit Exams

Although NCLB sanctions are aimed at schools, many states and districts are responding by holding students accountable for their own success or failure to meet performance targets. Proponents of that strategy hold that greater accountability gives students an immediate stake in their own progress and thus increases incentives for students to perform well. Among the major strategies for ensuring student accountability are the retention of students who fail to pass standardized tests at key grade levels and the use of exit examinations as a prerequisite for high school graduation. Those strategies are not elements of NCLB, but many states and districts are employing them in an effort to include students themselves in the spectrum of parties to be held accountable for student performance, and the testing requirements of NCLB create the mechanisms that enable these sanctions.

In chapter 5, Robert M. Hauser, Carl B. Frederick, and Megan Andrew examine trends in grade retention from 1996 to 2005. They use census data to examine national trends and data from fourteen state education agencies to consider state-specific patterns. A major concern regarding the relation between NCLB and poverty is that if school districts hold back students who fail assessments, the proportion of students below age-appropriate grade levels is likely to increase, particularly among students from minority and low-income backgrounds who tend to have lower test scores. Based on a cogent review of the existing literature, Hauser and his colleagues note that retention is associated with higher dropout rates and has rarely been effective as a means of boosting achievement. Consequently, the question of whether NCLB has resulted in a rapid rise in retention rates takes on great urgency. What the authors find, however, is no spike in retention. In fact, retention rates peaked in 2001, the year before NCLB was implemented. The national data show a consistent increase in retention from 1996 to 2001, likely reflecting the standards-based reform movement, but that trend was consistent with an older pattern going back at least to the 1970s. Thus, whereas stan-

dards-based reform probably has resulted in more use of retention, the early implementation of NCLB has not accelerated that practice. However, both national and state data show increased rates of retention in the earliest years of schooling—those not covered by NCLB—and the available state data show substantial variation in retention rates from year to year.

Hauser and his colleagues offer a tantalizing finding concerning social background and retention under NCLB: from 2001 to 2003, the overall degree of inequality in retention declined. Does that mean that under NCLB, retention decisions have become more "objective"—that is, less dependent on social background and more dependent on test scores (to the degree that background and test scores are independent)? Regardless of the cause, it is not yet clear whether the reduction in inequality will be sustained. In 2004, the relation between retention and social background jumped back to its pre-NCLB level, but it declined again in 2005. Hence, the relation between NCLB, retention, and inequality remains an open question that demands continued close scrutiny.

The use of high-stakes tests for high school graduation is another incentive for students under standards-based reform. In chapter 6, Thomas S. Dee and Brian A. Jacob examine the impact of high school exit examinations on high school completion, college enrollment, employment, and earnings. Also using census data, they estimate state fixed-effects models for the period 1980–98. In addition, they draw on the Common Core of Data to estimate district fixed-effects models for one state, Minnesota, from 1993 to 2002. An important feature of Dee and Jacobs' study is that the authors distinguish between "minimum competency exams," basic skills tests that have long existed, and more rigorous exit examinations that reflect the more recent standards movement. They also distinguish effects for different population subgroups, an approach that is essential for understanding the implications of exams for inequality.

Dee and Jacob find that states that introduced high school exit examinations between 1980 and 1998 tended to reduce their rates of high school completion, particularly among African American students. In light of the relation between race and poverty in the United States, that finding implies that exit exams may help perpetuate poverty rather than provide a way out. The finding held for both minimum competency exams and more rigorous testing schemes. For Whites and Blacks, exit examinations were unrelated to college enrollment or employment, but for Hispanic females, exit exams led to higher rates of college enrollment and employment. The coefficient for exit exams on high school completion also was positive for Hispanics, though

nonsignificant. Consequently, while exit exams may have exacerbated Black-White inequality, they seem to have mitigated inequality between Hispanics and non-Hispanic Whites. Dee and Jacob further observe that while minimum competency examinations were unrelated to earnings, more rigorous assessments contributed to higher weekly wages for employed African Americans but lower wages for Whites and Hispanics. That finding highlights the importance of considering labor market as well as education outcomes and hints that members of disadvantaged groups who succeed can benefit from high-stakes testing but that those who fail may fall further behind.

Findings on Sanctions for Schools: Supplemental Educational Services and School Choice

One way that NCLB differs from past standards-based reforms is that it includes not only standards but also serious sanctions intended to create better learning opportunities for students whose schools fail to reach their achievement targets. One of the primary sanctions is the mandate to provide "supplemental educational services" (SES)—that is, free after-school tutoring—to low-income students in schools that fail to meet their AYP targets for three years in a row. That requirement is expected to narrow the poverty gap by elevating the achievement levels of low-achieving, low-income students. In chapter 7, George Farkas and Rachel E. Durham examine what is known about SES and whether it is likely to achieve its goals. Drawing on a combination of past research, new government-sponsored case studies, and their own investigations, they address several questions about tutoring programs and whether they are likely to reduce the poverty gap.

Across the country, fewer students are receiving SES than might be expected on the basis of the large number of schools that have failed to achieve their AYP goals. The best recent estimate is that only 20 percent of eligible students are receiving supplemental services. Moreover, while the effects of SES are not being tested directly, there is good reason for concern about the benefits of this costly strategy for gap-closing, at least as it currently is implemented. Past research has provided convincing evidence of the benefits of one-to-one tutoring (one student per tutor), but tutoring under NCLB occurs mainly in small groups, for which there is much less evidence of effectiveness. Moreover, a nationwide randomized trial indicates no achievement benefits of after-school programs; at best, such programs are effective for elementary student achievement only when they are of high quality, as often is not the case. Farkas and Durham conclude that as currently designed, supplemental educational services are likely to have little

impact on achievement overall or on the poverty gap. To attain better results, tutoring needs to occur in smaller groups. Programs also must be more accessible, and students need to attend more regularly. Competition among providers might stimulate improvement, but only if sound information becomes available about which providers and types of programs are effective and only if parents act on that information at the local level.

School choice is yet another sanction that is mandated by NCLB and envisioned as a remedy for low test scores. Students whose schools fail to meet their AYP targets for two years in a row have the option to transfer to another public school, including a charter school, that has not failed AYP. The benefits of school choice for reducing inequality presumably work in two ways. First, obviously, it is supposed to give low-achieving students the chance to move to better schools. Second, it is expected to create incentives for districts and schools to elevate the quality of schools that are vulnerable to losing their students under school choice, thereby improving quality for all students and especially for those in schools with heavily disadvantaged populations.

In chapter 8, Paul T. Hill takes a hard look at the prospects of school choice under NCLB for reducing the poverty gap. He identifies both possible risks and potential benefits and gleans from existing evidence new ideas for how the risks can be managed and the benefits attained. Even though much has been written about the effects of both public and private school choice, Hill shows that there is little basis for firm conclusions. Implementation of NCLB is too recent to permit direct findings; moreover, very few students have exercised choice under NCLB. Even when previous work is considered, the evidence is so weak that it is difficult to predict what will happen under NCLB. That weakness is especially evident in the case of charter schools, which have generated a flurry of studies but few solid findings. Hill's strongest conclusion is that choice needs to be subjected to rigorous research, not only about whether the effects are positive or negative but also about the conditions under which school choice under NCLB can raise achievement and reduce inequality—and for whom. The evidence so far suggests that choice effects will be contingent on implementation, and understanding the contingencies will be the key to predicting the effects of choice on inequality.

The last section of the book looks toward the future of standards-based reform by drawing lessons from recent history and current practice. Understanding the politics of NCLB is one key to anticipating its future course. As Tom Loveless explains in chapter 9, NCLB was distinctive, in an era of polarized politics, in the level of bipartisan support that it initially enjoyed. While

support from across the political spectrum is still evident, so is criticism from all sides. According to public opinion polls, general support for NCLB has eroded to the point that now supporters and opponents are about evenly balanced. Still, the lines of support for and opposition to NCLB continue to confound convention. For example, supporters are found more often among Republicans and opponents among Democrats, but African Americans and Hispanics tend to be the most supportive demographic subgroups. Moreover, middle-income groups tend to be more supportive of NCLB than either low- or high-income respondents, who express more skepticism. Nor do views of NCLB at the state level conform easily to well-known patterns. Perhaps the most provocative of Loveless's findings concerns the relation between state performance on NAEP and state resistance to NCLB. The most resistant states tend to be those in the middle of the performance range; both high- and low-performing states tend to be less defiant. Moreover, states with the largest Black-White achievement gaps (and states with the smallest black populations) tend to be among the most resistant to NCLB. Whether NCLB can survive, Loveless reasons, depends on whether the members of the original coalition that supported the law—which Loveless characterizes as representing "conservative ideas . . . wrapped in liberal clothing"—continue to see advances from their various perspectives.

In chapter 10, the concluding chapter, Andrew C. Porter takes stock of the evidence presented in the volume on past practices of standards-based reform and offers insights for the reauthorization of NCLB. Porter begins, however, by focusing on aspects of NCLB that were not touched on elsewhere; in particular, he demonstrates dramatic differences among states in implementing NCLB. In light of Foorman and her colleagues' assertion in chapter 2 that the effects of standards-based reform depend on its implementation, the wide variation among states raises questions about NCLB's chances for success but also offers prospects for research, since variation is required in order to identify policy effects.

Porter's advice is drawn from a combination of evidence to date and a theory about how standards-based reform can be effective. According to that theory, an accountability system will succeed only if it sets a good target and if it is symmetric and fair. Using his theory, Porter points out a number of key strengths of NCLB, as well as some ways it might be improved. For example, he argues that assessment targets could be strengthened by attending to scores throughout the range of achievement instead of focusing on a particular proficiency target. Assessment targets could also be improved by addressing student mobility and by considering alignment between content

standards and assessments. The fairness of NCLB could also be strengthened; at present, neither supplemental services, nor school choice, nor highly qualified teachers are provided consistently or well. Without these resources, schools lack much chance of raising the achievement levels of their students, particularly students from the most disadvantaged backgrounds.

Forecast for Improvement

Will NCLB help reduce the poverty gap? While our crystal ball remains cloudy, the evidence and analyses in this book portend as much rain as sunshine. On the one hand, many of the greatest fears commonly expressed by opponents of NCLB have not materialized. Good teachers have not been driven away from the classroom by standards-based reforms; retention has not shot upward; and while high school graduation may be threatened by exit examinations, there is at least a hint that such exams may contribute to higher wages for African Americans in states that introduce them. NCLB clearly has brought attention to the inequality in achievement faced by students from racial, ethnic, and linguistic minorities; by those with economic disadvantages; and by those with disabilities. Moreover, there are some signs of progress related to standards-based reform, even if it is too early to judge NCLB. Certain aspects of teacher quality have improved, and some students are no doubt benefiting from new opportunities such as free tutoring and the chance to transfer to a different school.

On the other hand, the positive developments fall far short of the degree of change envisioned under NCLB. As Foorman and her colleagues point out, identifying inequalities, setting standards, and developing strategies are only the first step toward closing achievement gaps. When it comes to implementation, much work needs to be done, and some trends are troubling. While retention rates have not spiked, they rose steadily during the standards-based reform period, to the disadvantage of African American, Hispanic, and low-income students in particular. The few benefits to teacher quality seem to be related to teacher accountability rather than school accountability, which is the target of NCLB; moreover, those benefits accrued to schools with advantaged populations at least as often as to those in disadvantaged environments, so little reduction of inequality can be seen. Use of the school choice and tutoring options, two key strategies for improvement, has occurred at such low rates and with such inconsistent and poorly monitored quality control that even the relatively few students who have participated may not have obtained the expected benefits. Nor, as Porter

notes, have the teacher quality provisions of NCLB been fulfilled, so the modest improvements identified by Phillips and Flashman and by Desimone, Smith, and Frisvold probably are not widespread.

Most of the debate and commentary that one hears and reads about NCLB has to do with its system of test-based accountability, as states, districts, and schools struggle to cope with seemingly unrealistic expectations of how fast test scores should rise. The findings in this book, however, suggest that an even greater problem with NCLB may lie in the inadequate implementation of strategies to respond to test score gaps. With attention focused on achievement inequality and with NCLB sanctions providing a strong motivation for change, a unique window of opportunity now exists to bring new resources and strategies to bear on the education of poor children.

Can NCLB's implementation be improved, and if so, will improvement mitigate achievement inequalities? That is the question that future work must address. At this point, four comments are in order. First, according to the findings in this book, the strategies promoted by NCLB remain promising. That is, improving teacher quality, offering choice and supplemental services, and promoting evidence-based practice have not been ruled out as strategies for addressing the poverty gap. They remain largely untested, but available evidence does not reject them. Second, in contrast, the strategy of extending accountability to students is dubious. Both retention and high school exit examinations have negative consequences for at least some students. It falls on proponents of these strategies to demonstrate their benefits before they can justifiably be extended. Tellingly, they are *not* part of NCLB; rather, they reflect the response of many states and districts to NCLB requirements. In this instance, NCLB seems to have gotten it right: sanctions directed at students are not likely to reduce inequality. Third, more resources need to be devoted to extending supplemental services and to studying its effects. That is not only an implementation challenge; it is also a major research task. Little is now known about even the most basic questions concerning tutoring under NCLB, such as whether participants are coming closer to proficiency over time. Fourth, more rigorous research on a wider variety of educational practices is needed, and the research must be designed explicitly to identify effective instruction for disadvantaged children. It is well to call for evidence-based practice, but as long as the evidence remains in short supply, the call will go unanswered.