Fulfilling the Promise of Educational Accountability

Widely regarded as the model for NCLB, the Texas education reforms have had their share of unintended consequences. The authors argue that examining those problems will suggest the changes that need to be made in both reform agendas.

BY SARAH W. NELSON, MARLA W. McGHEE, LIONEL R. MENO, AND CHARLES L. SLATER

WHEN No Child Left Behind (NCLB) was signed into law, the President and secretary of education promised sweeping reform of the American education system. In the five years since the law took effect, U.S. public schools have, indeed, seen remarkable change. However, not all of the changes have been well received. Policy makers, scholars, and practitioners alike are increasingly voicing concerns about this comprehensive piece of legislation.

In fact, legislatures in 47 states have taken action to mitigate the effects of NCLB.[1] These legislative responses range from petitioning the U.S. Department of Education for waivers to refusing to comply with some or all aspects of the law. Indeed, the Connecticut legislature went so far as to file a lawsuit challenging the legality of NCLB and the use of the policy to guide distribution of federal entitlement funds.[2] Even Texas, widely considered the birthplace of NCLB,[3] showed resistance: its commissioner of education refused to adhere to the NCLB rule that limits to 1% the number of students who may be exempted from testing because of learning disabilities. Rather than comply with this mandate and label schools that exceeded the limit "low performing," the Texas commissioner gave acceptable accountability ratings to more than 900 schools with exemptions greater than 1%.[4]

In response to this growing resistance, political and professional groups are calling on Congress to make changes to NCLB policy. Recently, a special task force of the National Conference of State Legislatures issued a comprehensive report documenting the findings of a 10-month study of the effects of NCLB. The report includes recommendations for significant changes to the law.[5] Similarly, the National Association of Elementary School Principals and the National Association of Secondary School Principals issued a joint statement indicating support for the goals and intent of NCLB but blasting the policy for being an unfunded mandate that is "unworkable" in its present form.[6]

As scholars and former public school practitioners, we, too, believe that modification of NCLB is warrant-

SARAH W. NELSON is an assistant professor in the Department of Education and Community Leadership at Texas State University-San Marcos, where MARLA W. McGHEE is an associate professor. LIONEL R. MENO is dean of the College of Education at San Diego State University, San Diego, Calif. CHARLES L. SLATER is a professor of educational administration at California State University-Long Beach.

ed. However, our call for change stems not so much from what has occurred in the five years under NCLB as from what we have learned in more than 15 years of work within the Texas accountability system. As the state whose reform effort served as the model for NCLB, Texas has much to offer in the debate about school accountability. In particular, examining policies and practices that have worked well in Texas along with those that have proved problematic may shed light on necessary changes for NCLB.

What we offer here are the history and lessons of accountability in Texas from the perspective of four individuals who worked within the system from the beginning: 1) the former Texas education commissioner who was one of the architects of the initial reform movement, 2) a former school district superintendent, 3) a former school principal, and 4) a former teacher. When we're asked, as we often are, whether Texas education reform has turned out as we expected, our consensus answer is, "Yes and no."

The Vision

In the early 1990s, Texas set out on a journey of education reform. Led by the state commissioner, this effort was noteworthy in that it challenged two long-held assumptions about education reform. First, whereas all education reforms naturally aim to improve student achievement, the Texas model was unique in focusing on improvement for *each student* rather than aggregate groups of students. That is, the Texas model was built on the notion that every student deserves to be well educated. Consequently, the state has a responsibility to provide an education system that gives each student, regardless of where he or she attends school, a high-quality education.

While one could argue that focusing on each student as opposed to the aggregate is a matter of semantics, given that improvement in the aggregate requires improvement by individual students, the developers of the Texas model did not see it this way. A focus on individual students was essential, they asserted, because improvement in the aggregate requires improvement only by some, not by all, students. Since education systems have traditionally served some, but not all, students well, a focus on the aggregate level would probably not result in meaningful change. An explicit focus on each student was necessary.

In addition, the Texas reform effort was unique in that the effectiveness of districts and schools was defined in terms of outcomes rather than inputs. Prior to the development of the Texas model, the quality of education in a district or school was measured by factors related to what happened *before* any learning occurred (e.g., per-pupil expenditures, number of available textbooks, correctly completed paperwork). The assumption that inputs equaled quality led to reform efforts aimed at frontloading the system as a means of improving education. While inputs are certainly significant factors in the quality of education, the Texas model considered outcomes as the most important measure of quality.

These two ideas — focusing on individual students and measuring outcomes — laid the groundwork for educational equity to become the focal point of the reform movement in Texas. While the rest of the nation was just beginning to consider the consequences of the achievement gap, Texas was in the midst of developing a reform model that would systemically alter the inherently inequitable nature of its public schools.

Implementation

Comprehensive reform in Texas rested on four critical components: 1) declaring what should be learned (curriculum), 2) measuring what is learned (assessment), 3) creating a system of public reporting and accountability, and 4) doing what is necessary to improve student learning. In many respects, the components of this system have improved public education in Texas. At the same time, the implementation of these components has generated unintended consequences that have negatively affected Texas schools. The challenge, then, is to modify the system in such a way as to maintain the improvements while mitigating the negative effects. To do so requires an understanding of what has worked and what has not.

WHAT WORKED?

Curriculum. The first statewide curriculum, referred to as the Essential Elements (EEs), was installed during the 1984-85 academic year. Before the advent of the EEs, the responsibility of curriculum development fell to individual school districts. In some instances teachers were expected to teach courses for which no curriculum had been established. In the absence of guiding documents, it was not uncommon for the selected textbook or instructor-generated materials to fill the gap. Course expectations and content varied greatly across the state and depended on the classroom, school, or district. The development and implementation of the EEs helped to address these curricular irregularities and ensured that each Texas student had access to a well-

rounded curriculum that included not only language arts, math, science, and social studies, but also other disciplines, such as theater arts, physical education, art and music education, and languages other than English.

As part of the reform, the state commissioner initiated an update of the curriculum to better address the needs of Texas students and to reflect the high-tech surge sweeping through the state and nation. Whereas the EEs were a description of what students would have the opportunity to learn and were often couched in terms of having students *recognize*, *understand,* and *explore* various curricular concepts, many thought it was time to declare what students would be able to do as a result of their learning[7] and to have students engage in more rigorous tasks such as *using*, *analyzing*, and *evaluating*.

To develop this more demanding curriculum, the Texas Education Agency solicited nominations of teachers, instructional supervisors, campus administrators, professors, business representatives, and parents to serve on curriculum writing teams. These teams, one for each subject area, developed draft documents, which were vetted by public school educators as well as by the general public and national content-area experts. The result was a comprehensive curriculum that specified what Texas students were expected to learn in order to be prepared to meet the demands of an increasingly complex and diverse world.

Assessment. To measure schools' and districts' progress toward meeting the goal of having every student master the state curriculum, the Texas reform model called for the development of a state assessment system. The idea of mandatory testing of students was not new in Texas. A statewide testing system had been in place since the mid-1980s. However, these early tests, known as the Texas Assessment of Basic Skills (TABS) and then the Texas Education Assessment of Minimum Skills (TEAMS) were, as the names imply, designed to measure mastery of minimum skills. Under the new reform model, the aim of assessment shifted to measuring mastery of higher-level skills.[8]

In addition, in keeping with the focus on achievement for each student, the new assessments, known as the Texas Assessment of Academic Skills (TAAS), were criterion-referenced rather than norm-referenced. Moreover, the TAAS was not purchased from a commercial test publisher but was developed under the direction of the Texas Education Agency to specifically align with the state curriculum and to be diagnostic in nature so that teachers could use the results to guide instruction. Since the tests were criterion-referenced and were based on the state curriculum, the implicit assumption was that every student should master the material on the test. In this way, the assessment system became not only a measure of what students had learned successfully in the past, but also an indicator of what they still needed to learn.

Public reporting. In conjunction with the new assessments, the Texas reform model included a performance monitoring system intended to measure educational progress over time. A key feature of this monitoring system, referred to as the Academic Excellence Indicator System (AEIS), was annual campus-, district-, and state-level reporting that allowed teachers, parents, and members of the larger community to "see" what students were learning or not learning. AEIS reports were distributed annually to school communities during mandatory public meetings and were made available to the general public via the state department's website.

Like other aspects of the reform model, public reporting was not an entirely new idea. Under the previous assessment system (TABS), annual reports of test scores were issued for each school and district. What distinguished the AEIS reporting system was the comprehensiveness of its data and the wide availability of those data to the public. AEIS reports included more than just test scores; they also featured data related to student attendance, dropouts, instructional program participation, student and staff demographics, and fiscal allocations and expenditures.

For the first time, every teacher, principal, and superintendent in the state had access to a consistent set of student performance data. Through electronic databases, this dataset could be queried to produce both individual and comparative performance reports on students, teachers, campuses, districts, and the state. Moreover, the data were reported by subject, grade level, and demographic group. In fact, the data were so easily accessible that analyzing test data in a variety of configurations became routine practice in most Texas school districts. Consequently, Texas educators tended to be highly informed about individual student performance and often used this information to make such instructional decisions as which areas of the curriculum needed greater emphasis and which students were eligible for additional support.

Of all the information available through the AEIS reports, the disaggregation of student test scores by ethnic and economic groups proved to be the most important. When the data were broken down in this manner, inequities within the education system were revealed. Test scores for students of color and students living in poverty consistently trailed those of white, middle-class students. While educators had long known that some

students were not being served well by the schools, the public disclosure of clear evidence of disparate outcomes made addressing the issue unavoidable. In this way, the architects of the Texas reform model created a transparent system of public education. Through this transparency, the citizens of Texas could, in theory, hold schools and districts accountable for educating all the state's children. As a result, educational equity became a focal point of discussion among practitioners and policy makers in Texas.

Support systems. As expectations for accountability and performance increased, steps were taken to expand the related support systems. For example, the legislature codified specific guidelines for student/teacher ratios in the early elementary grades. Through grade 4, no more than 22 students could be assigned to a classroom. This requirement was accompanied by other examples that emphasized the importance of early learning experiences, such as the implementation of full-day kindergarten programs for all and of half-day prekindergarten programs for English-language learners and children living in poverty.[9]

In addition to providing support for students, the Texas reform model called for improving teacher quality through professional development.[10] In the early 1990s, the state commissioner advocated increasing the number of contract days for Texas teachers to include two to three weeks a year of paid time for professional development. At the same time, the commissioner declared a moratorium on beginning-of-year inservice training that did not promote learning experiences specifically tied to a school's student learning goals. This generated a move away from large, whole-group, one-size-fits-all professional development sessions and toward customized learning pursuits tailored to the targeted needs of teachers and students at each school.

UNINTENDED OUTCOMES

Not surprisingly, with its sound foundation and focus on excellence and equity, the Texas reform model produced positive results almost instantly. Test scores seemed to rise overnight, and schools that had a history of failure were finding ways to succeed. However, just as rapidly, unintended consequences began to emerge.

Curriculum. Development of a new curriculum aimed at increasing the level of academic performance generated considerable controversy along the way. A draft of this new curriculum, Texas Essential Knowledge and Skills or TEKS, was released for review in August of 1996. Many who interacted with the 2,000-page draft complained that the objectives were written in ways that were not easily *testable* — an indication that the assessment system, designed to measure what students *had* learned, was beginning to drive what students *would* learn.

The strong influence of the assessment system was also seen in the organization of the new curriculum. Whereas the previous statewide curriculum gave equal value to all subjects, the TEKS was presented in two separate parts, one known as the "foundation curriculum" and the other called the "enrichment curriculum." Curricula for fine arts, health and physical education, languages other than English, and technology applications were no longer considered part of the central curriculum and were relegated to an "enrichment" status. Although districts were still required to offer the enrichment curriculum, the emphasis had clearly moved away from *teaching* a well-rounded curriculum to testing specific subjects.

The influence of testing also affected the curriculum in other ways. Curriculum narrowing has been widely documented. While language arts and math were regularly tested on state assessments, subjects such as science and social studies, until only recently, were not. Teachers reported instances in which district and campus leaders encouraged the teaching of the tested subjects over those not assessed.[11] This practice seemed to be especially prevalent in schools serving predominately students of color or poverty.[12]

In addition, new classes of subject matter have sprung up as a result of accountability testing. To prepare them for the yearly Texas Assessments of Knowledge and Skills (TAKS), students are made to take "TAKS Math" or "TAKS English." The students "know these are not real courses; the content is watered down and fragmented, arranged to fit the upcoming test rather than to bridge students into college-preparatory levels of these subjects."[13] Again, the impact has been more prominent for students who have been historically marginalized by school systems, and we are beginning to see the effect of this "substitute" curriculum over many years of a child's school experience. Year after year in some schools, children of color have been subjected to test-prep activities and materials in lieu of the regular curriculum, with the result that academic gaps in nontested subject matter have been exacerbated rather than reduced.

Assessment. Under the original vision of a comprehensive accountability system in Texas, the assessment system would have culminated in a capstone experience for high school students, such as a senior project. To ensure that students reached this mark, the system

was originally designed to use a combination of performance-based and multiple-choice testing. The introduction of performance-based tasks was intended as a first step in creating a progressive assessment system aimed at measuring synthesis of learning rather than mastery of discrete knowledge. The system design also called for assessing the full curriculum through sampling rather than assessing every student on a narrowed curriculum.

This vision of the assessment system, however, never came to fruition. After just the second field test, performance assessment was abandoned in favor of less costly and easier-to-administer multiple-choice tests. Moreover, the scope of the curriculum included in the assessments was narrowed to just a few subjects, and the notion of sampling was dropped.

This change of course facilitated a quick expansion of testing. In 1990, when the TAAS was administered for the first time, only general education students in grades 3, 5, 7, 9, and 11 were tested and then only in math, reading, and writing. In 1994, TAAS was extended to include testing at grades 4 and 8. In 1995, social studies and science tests were added to grade 8-testing, and end-of-course assessments were added for algebra I and biology. In 1998, end-of-course examinations for English II and U.S. history were added. During this same year, a Spanish version of the test was made available for the first time so that bilingual students in grades 3-5 could be included in the system.[14] Similarly, in 2000, an alternative version of the test was developed so that qualified special education students could be tested. In fact, the Texas legislature or state board of education has altered the testing system nearly every year since 1990.

Predictably, an assessment system of this size generates an enormous amount of data. While the practice of analyzing test data increased educators' awareness of student performance and allowed them to adjust the curriculum, the focus on test data also created an environment in which scores from standardized tests came to be viewed as the primary — and sometimes the only — valid measure of performance.

Placing such a high value on accountability test scores has, in turn, encouraged educators to periodically administer "practice" tests in an attempt to determine whether students are on track to pass the "real" test. Across the state, schools and districts sanctioned test preparation as a legitimate aspect of schooling by devoting time and financial resources to practice tests. In some districts, students take practice tests and attend test-prep classes as part of their daily routine. In this way, measuring what has been learned has become as important as teaching itself.[15]

Public reporting. In the early 1990s, a rating scale was added to the AEIS. This four-level scale rated schools and districts on how well each was meeting state standards on key AEIS performance indicators. At a minimum, schools and districts were expected to meet standards for a rating of "acceptable." Schools and districts receiving the lowest ratings ("academically unacceptable" or "low-performing") were subject to sanctions.

The primary benefit of this rating system was the ease with which parents and community members could evaluate schools and districts. It was widely assumed that "exemplary" ratings connoted high-quality schools and districts, while "low-performing" or "unacceptable" ratings marked ineffective schools and districts. However, this simplification of schools' and districts' effectiveness proved problematic. Because the ratings were based on some, not all, AEIS data, the ratings often masked inequitable practices. Some districts were able to find loopholes in the system that allowed them to maintain high ratings without actually fulfilling the obligation to teach all students.

For example, from 1994, when the ratings were introduced, through 2002, the inclusion of special education students in accountability testing was not required. Some schools protected their high accountability ratings by inappropriately placing struggling students in special education programs and exempting them from the test, effectively denying these students access to the regular curriculum. More troubling, such practices disproportionately affected students of color and students of poverty. There were, for example, "exemplary" schools that exempted 50% or more of their African American students.[16] Similarly, the sanctions associated with low performance ratings disproportionately affected schools serving high populations of children of color and children of poverty. In 2006, 11 schools were rated "low-performing" for the third year in a row and were thus subject to reconstitution. All 11 are high-minority, high-poverty schools.[17]

Moreover, while accountability ratings were originally intended to indicate the degree to which schools were meeting state standards, the ratings have come to characterize the people associated with the schools. Principals and teachers who work in "recognized" and "exemplary" schools are typically viewed as being highly competent; those who work in "low-performing" schools are often seen as ineffective and are blamed for the schools' low ratings. This assumption discourages many educators from working in lower-rated schools, which typically are schools serving large numbers of low-income and

minority students. While staffing such schools with highly qualified educators has long been a challenge,[18] the ratings associated with the accountability system have made it that much more difficult.

More troubling still, this practice of characterization according to performance on an accountability test has been extended to students. As analyzing test data became common practice, so too did the practice of sorting students by their likelihood of passing the test. Those students whose past achievement indicates that they are likely to pass the test without intervention are often given access to an enriched curriculum. Those students whose achievement is near the passing mark have come to be known as "bubble kids" because they are viewed as being on the brink of passing the test. Such students are often given intense intervention to ensure they reach the mark. Students whose previous achievement indicates they are highly unlikely to pass the test even with intense intervention are viewed as too far behind to warrant attention. While most such students are given some level of assistance, the focus is clearly on those who are seen as having the greatest chance of passing. Thus some students have essentially been written off as early as third grade. An article by Joshua Benton of the *Dallas Morning News* shows how the "bubble kid" concept works:

> In the fall, teachers gave students a sample TAKS test. Based on the results, students were divided into three groups: passers at the top, remedial kids at the bottom, and bubble kids in between. The bubble kids are the ones whose scores put them just below the state's passing standards . . . the ones who, with a coordinated effort, can be pushed over the passing bar. . . . So how did the educators at this particular school react? By pouring all the resources they could into the bubble kids. The bubble kids get special sessions with the school's reading specialist. The bubble kids get after-school and Saturday tutoring. The bubble kids get small-group attention in class. The bubble kids get extra reading time with librarians and the P.E. teacher. . . . Teachers aren't stupid. They realize they're going to be judged on how many of their kids pass — not how much improvement they can squeeze out of their weakest kids. So they go after the low-hanging fruit: the bubble kids. . . . I've talked to dozens of teachers who do some version of the same practice. Principals call it being "data-driven." I call it an excuse to ignore the weak.[19]

Support systems. While the initial concepts related to the support systems had great potential at the time they were introduced and implemented, few have remained in place in their original form or been sufficiently supported to achieve their fundamental promise. Take, for example, professional development. Except in a limited number of model partnership schools and a few districts that added extra days on their own, the additional contract days for professional development never materialized. Moreover, because of the press for positive accountability results, a new educational vocabulary and line of professional learning emerged, which shifted the focus of professional development away from improving teaching and learning and toward raising test scores. In this new genre of professional development, paid consultants crisscrossed the state urging teachers to focus their teaching time and strategies on the "bubble kids" — often at the expense of other students who were further behind academically. Other common staff development sessions focused only on how to beat the test, arming teachers with ways to help students use test-taking strategies to score well on their accountability assessments without regard for authentic learning experiences.

Midcourse Corrections: Recapturing the Vision

Over the past 15 years we have learned much about accountability systems — their positive outcomes and their unfortunate and unintended consequences. Just as Texas has led the nation in establishing and practicing educational accountability, the state should also offer critical lessons about the pitfalls inherent in operationalizing a "Texas Style" approach to large-scale reform. We believe that the steps we outline below are necessary in order to refocus improvement initiatives in Texas (and across the nation) so that they may finally accomplish the promise that was the original vision of the accountability program.

STANDARDS, NOT STANDARDIZATION

While it is one thing to have high expectations and standards of performance for all youngsters, it is quite another to expect all youngsters to achieve that high degree of performance at precisely the same time. If we have learned anything at all from interacting with children over time, it is that there are undeniable differences among them. Students come to school each day with a variety of interests, gifts, and strengths. Some learn with seemingly little instruction while others require additional time or an alternative route. Rather than continue to impose unrealistic deadlines, time lines, and consequences that fail many learners, we should instead develop processes that engage parents, educators, and stu-

dents in meaningfully crafting pathways to support learning and achievement.

This is not to say we should expect less from some students. Quite to the contrary, we should expect that all students meet rigorous standards, but we should allow them to follow different pathways. Valuing assets rather than dwelling on deficits could revolutionize attitudes about the richly diverse population of students attending our schools and provide the flexibility needed to see that all children achieve at high levels. To realize such a lofty vision, we must distinguish between standards and standardization. The goal must be to help students meet rigorous curricular standards, not to ensure that every student has a standard educational experience.

FAIR AND COMPREHENSIVE ASSESSMENT

Although the current statewide curriculum is clearly rigorous and comprehensive, the state examines relatively few aspects of that curriculum. In addition, testing is conducted in a narrowly focused, multiple-choice, nonintegrated format. If we sincerely believe the best proof of learning happens when students are asked to apply their knowledge and skills in real-world situations, then that is how we should be assessing them. To act as if this type of assessment is not feasible on a large scale is, in fact, disingenuous. For years the National Assessment of Educational Progress (NAEP) has successfully assessed learners in an open-construction format, where students use what they have learned to create something new. This approach may be more time-consuming, but it is surely a more authentic and appropriate gauge of student learning than the current practice of single-subject, objective testing.

Moreover, the notion of using a single indicator to make such high-stakes decisions as promotion or graduation is, at best, misguided. Even those who have dedicated their careers to creating valid and reliable tests assert that no one test could ever be sufficient to adequately measure the totality of student learning.[20] Moving to a multiple-indicator system for students and schools would be a more accurate and just way to assess progress and would tend to reduce some of the most harmful unintended outcomes, such as curriculum narrowing.

ADEQUATE SUPPORT SYSTEMS

Like many states, Texas has struggled in recent years with the issue of public school finance. In a state with a rapidly increasing school-age population and one that prides itself on a low per-capita tax rate, providing funding to maintain a solid educational infrastructure is a contentious and unsettled topic. And while we do not suggest this issue is easily addressed, we do submit that, to uphold the integrity of the educational accountability system, adequate funding to maintain support systems must be afforded to public schools.

Such support systems as low teacher/pupil ratios, early childhood education, and high-quality professional development programs have proved instrumental in creating the impressive gains in academic achievement that have been so widely reported. If such gains are to continue, we must not only maintain but expand such support systems. To do otherwise is to undercut the promise of accountability and to ignore the high costs of underfunding education — a growing achievement gap, a less-educated citizenry, and higher long-term costs. In a time when many are arguing that public schools are more than adequately funded given their mediocre performance, we must be courageous enough to counter such arguments with evidence and to demand adequate support for what works in education.

FOCUS ON THE SYSTEM, NOT INDIVIDUALS

While a guiding aim of the Texas reform model was

to focus on high-quality education for each student, the designers of the reform believed that the responsibility for ensuring a high-quality education rested with the *system* — the state, the public school districts, school sites, and the classrooms. The reformers reasoned that only by focusing at the systems level could we eradicate the systemic inequities that had plagued education in Texas. With a systems approach in mind, the reformers planned to use a sampling method in assessment, much like the system pioneered by the NAEP. Statistically significant sampling would generate sufficient data to measure whether the state and districts were serving all students well and would allow the use of authentic testing formats that might be cost-prohibitive if given to every student.

However, the sampling approach was never employed. A desire to measure the learning of individual students led to the institution of annual testing for every student in almost every grade and subject. While this approach has generated valuable data related to individual performance, it has also shifted the focus of the reform from system accountability to individual accountability. When test scores are not high enough, someone — student, parent, teacher, or principal — is presumed to have failed for lack of effort or skill. As a result, interventions are aimed at improving individuals rather than improving the system. Consequently, in spite of 15 years of accountability and improvement by individual students, teachers, and schools, the systemic inequities and the problems that led to the initial reform (e.g., inadequate and inequitable funding, low achievement, low teacher quality, dropouts, low college entrance rates) remain.

Therefore, while we believe that individuals should be held accountable and that individual achievement data can be useful, we also believe that equal effort must be made to address systemic inequities if we are to reach our goal of a high-quality education for every child. It is time to pull back the lens of accountability and focus on the *system*.

We strongly encourage educators to come to the policy-making table. Much of the accountability policy we operate under today was written without the benefit of input from those most affected by it. Educators can no longer afford to sit idly while laws are being passed; policy makers can act only on the information they have. We must help lawmakers understand the connection between education and public policy. We must help them recognize and honor the importance of educated citizens to our democracy, our society, and our economy and to understand that accountability to children and youths extends beyond the schoolhouse doors. Educators, while not abdicating their responsibility to educate each child to high standards regardless of their background, must help policy makers and the public understand that, when American schoolchildren have adequate housing and health care; are provided appropriate preschool, after-school, and summer programs; and have parents who earn a living wage, schools will work better, students will flourish as learners and as leaders, and the promise of comprehensive accountability will be fulfilled.

---

1. Communities for Quality Education, *State and Congressional Summaries*, available at www.qualityednow.org/summaries.
2. Robert A. Frahm, "'No-Child' Case Draws Nationwide Interest," *Hartford Courant*, 15 April 2005, p. B-1.
3. Achieve, Inc., *Aiming Higher: Meeting the Challenges of Education Reform in Texas*, June 2002, available at www.achieve.org/node/277.
4. Jason Spencer, "State Defies U.S. Rules on Grading Schools," *Houston Chronicle*, 26 February 2005.
5. National Conference of State Legislatures, "State Legislators Offer Formula for Improving No Child Left Behind," *NCSL News*, 23 February 2005, available at www.ncsl.org/programs/press/2005/pr050223.htm.
6. Vincent L. Ferrandino and Gerald N. Tirozzi, "Improving NCLB: A Principal's Perspective," 15 March 2004, available at www.naesp.org/ContentLoad.do?contentId=1215.
7. "A Collaborative Effort of the Texas Education Agency," *Technical Digest for the Academic Year 2004-2005*, Pearson Educational Measurement, Harcourt Educational Measurement, and Beck Evaluation and Testing Associates, Inc., 2005.
8. Keith Cruse and Jon Twing, "The History of Statewide Achievement Testing in Texas," *Applied Measurement in Education*, vol. 13, 2000, pp. 327-31.
9. Achieve, Inc., op. cit.
10. Ibid.
11. Linda McNeil, *Contradictions of School Reform: Educational Costs of Standardized Testing* (New York: Routledge, 2000).
12. Angela Valenzuela, "The Significance of the TAAS Test for Mexican Immigrants and Mexican American Adolescents: A Case Study," *Hispanic Journal of Behavioral Sciences*, vol. 22, 2000, pp. 524-39.
13. McNeil, p. 247.
14. Cruse and Twing, op. cit.
15. Sarah W. Nelson and Marla W. McGhee, "Time Off Task: How Test Preparation Is Siphoning Instructional Time for Students of Color and Students of Poverty," paper presented at the annual meeting of the University Council for Educational Administration, Kansas City, Mo., 2004.
16. See Academic Excellence Indicator System (AEIS) 2002, report for campus #227901146, available at the Texas Education Agency, www.tea.state.tx.us/perfreport/aeis/2002/index.html.
17. See Texas Education Agency Accountability Ratings System, available at www.tea.state.tx.us/perfreport/account.
18. Linda Darling-Hammond, "Inequality and the Right to Learn: Access to Qualified Teachers in California's Public Schools," *Teachers College Record*, October 2004, pp. 1936-66.
19. Joshua Benton, "TAKS Push Not So Equal," *Dallas Morning News*, 19 September 2005, p. 1-B.
20. "Miles to Go Before We Sleep," *Activity*, Winter 2004, available on the ACT website at www.act.org/activity/winter2004/miles.html. K

File Name and Bibliographic Information

k0705nel.pdf

Sarah W. Nelson, Marla W. McGhee, Lionel Meno, and Charles L. Slater, Fulfilling the Promise of Educational Accountability, Phi Delta Kappan, Vol. 88, No. 09, May 2007, pp. 702-709.

Copyright Notice
Phi Delta Kappa International, Inc., holds copyright to this article, which may be reproduced or otherwise used only in accordance with U.S. law governing fair use. MULTIPLE copies, in print and electronic formats, may not be made or distributed without express permission from Phi Delta Kappa International, Inc. All rights reserved.

Note that photographs, artwork, advertising, and other elements to which Phi Delta Kappa does not hold copyright may have been removed from these pages.

Please fax permission requests to the attention of KAPPAN Permissions Editor at 812/339-0018 or e-mail permission requests to kappan@pdkintl.org.

For further information, contact:

Phi Delta Kappa International, Inc.
408 N. Union St.
P.O. Box 789
Bloomington, Indiana 47402-0789
812/339-1156 Phone
800/766-1156 Tollfree
812/339-0018 Fax

http://www.pdkintl.org