This project was funded by Carnegie Corporation of New York. The statements made and views expressed are solely the responsibility of the authors.

Copyright © 2007 by The Aspen Institute

The Aspen Institute
One Dupont Circle, NW
Washington, DC 20036-1133
Published in the United States of America
in 2006 by The Aspen Institute

All rights reserved
Printed in the United States of America
ISBN: 0-89843-470-X

1593/CP/BK
Table of Contents

Rapporteur’s Summary ............................................................. 1
Carol Copple

Education as the Cornerstone of Global Competitiveness:
How Does the U.S. Compare Internationally? ................................. 9
William A. Galston

No Child Left Behind: Revisiting Adequate Yearly Progress .................. 13
Barbara Byrd-Bennett

Measuring Proficiency With Status and Growth Models:
How Best to Maximize the Achievement of All Students? .................. 19
James Wyckoff

What Can a School of Education Do to Prepare High Quality
Teachers for Low-Performing Urban Schools? ............................... 27
Victoria Chou

Can the No Child Left Behind Act Solve the Problem
of Underqualified Teachers? ......................................................... 33
Richard Ingersoll

The Challenge of Special Populations to Accountability for All .............. 39
Martha Thurlow

Do We Need National Standards? .................................................. 45
Michael Cohen

Participant List ................................................................... 51

Conference Agenda .................................................................. 53
“No Child Left Behind: A Five Year Review” was the theme of the 2007 conference sponsored by The Aspen Institute Congressional Program, held February 20-25 in San Juan, Puerto Rico. The meeting was the fourteenth in a series that examines policy options for promoting the education and well-being of American youth, and thus the well-being of the nation. A bipartisan group of 17 Members of Congress participated, together with seven invited experts with relevant knowledge and experience. The Congressional Program series is not intended to yield a consensus statement of recommended policy directions. Rather, the aim is to help inform policymakers and facilitate the search for common ground on which effective American legislative policy must rest. The 2007 Congressional Program conference examined a range of issues and policy options with respect to achieving the goals of No Child Left Behind (NCLB).

Standards-based reform and NCLB accountability have succeeded in getting the attention of educators nationwide. Five years since the advent of NCLB, some achievement gains in school systems are evident. However, real progress is neither as robust nor as widespread as it might appear at first glance. While all state testing programs are able to point to achievement gains, there is enormous variation in the rigor of the various states’ assessments and how they define proficiency. The common metric provided by the National Assessment of Educational Progress (NAEP) reveals wide state-to-state differences and high proportions of students performing below the proficient level in reading and other key areas. The public schools may have begun to move in the right direction, but many states, districts, and schools have a long way to go in enabling all students to achieve proficiency.

Time is not on our side. Other nations are racing to compete with the United States, and the global, technology-intensive economy is evolving rapidly. As citizens and as workers, Americans must have the knowledge and skills to succeed in this fast-changing environment. As a nation, we need to reach higher and more universal levels of student proficiency as swiftly as possible.

The No Child Left Behind Act has given public schools in the U.S. a clear message that business as usual will not suffice. Proficiency for all students has become the new expectation. NCLB allows states to use their own standards and tests to determine proficiency, but to ensure that all students are progressing well they must report achievement data disaggregated by groups within the school population. Recognizing that to reach the goal of universal proficiency, students would need competent teachers, lawmakers also included in NCLB the requirement of a “highly qualified teacher” in every classroom by the 2005-06 school year. This provision, along with the rest of NCLB, is due for reexamination as the time for the reauthorization approaches.
A highly qualified teacher in every classroom—Are we there yet?

Teachers have a profound and lasting effect on student achievement, research clearly shows. Moreover, teacher impact is especially significant for low-income and minority students, who as a group have the greatest learning needs. NCLB affirmed teachers’ importance in the highly qualified teacher (HQT) requirement, giving states the right to define what “highly qualified” means but insisting that the HQT have at least a bachelor’s degree, full state certification, and demonstrated competency in the core academic subject area assigned. By and large, the states have met the first two requirements for the vast majority of teachers. Shortfalls still exist in places, but across the board the number of teachers lacking a degree and full certification comprise only a small proportion of the K-12 public teaching force.

For attaining the goal of a highly qualified teacher in every classroom, the main deficiency lies with the third criterion. Substantial numbers of teachers who otherwise meet the HQT definition are teaching outside the fields in which they are competent. Equity is also an issue. While high-poverty schools lag only slightly behind other schools in the formal qualifications of their faculty, they are far more likely to have teachers teaching out of field. Nor is there evidence that such assignments have diminished significantly since passage of NCLB. Analysis of school staffing data indicates little change in out-of-field teaching in the years since NCLB became a law.

The main reason for misassignment of teachers is not the teacher shortages in high-demand fields like science and mathematics; data indicate that out-of-field teaching is also prevalent in areas for which teachers are in surplus, such as English and social studies. Rather, principals report that they find assigning teachers to teach out of their fields is often more convenient, less expensive, and less time consuming than the alternatives.

If assignment decisions were shared with faculty, who must live with and be accountable for the consequences, out-of-field teaching would likely decrease. States and districts therefore could take action to encourage giving teachers a say in their assignments. Another option is for states to provide training and assistance to district and school administrators in how to meet organizational and budgetary demands without sacrificing students’ right to qualified teachers. The federal government could encourage these actions using sanctions, incentives, or both.

Although in NCLB Congress affirmed the importance of teacher quality and reducing the inequity in teacher distribution, states have not yet significantly narrowed the equity gap. In fact, many states seem to have chosen to ignore the legislative intent behind NCLB. Rather than make the changes required to close equity gaps, they have focused principally on compliance and have found many ways to “game the system.” Hence, federal leadership and direction in efforts such as raising teacher quality are critically important, especially in struggling schools. The reauthorization of No Child Left Behind presents opportunities to reduce the gaming of the system and provide tools and support for states, districts, and schools to make real strides.

Qualified and effective teachers

The highly qualified teacher criteria stated in NCLB constitute a legal definition and seem to provide a reasonable baseline for qualified teachers. But they are not a full prescription for classroom effectiveness. In reality, many well credentialed teachers lack the skills to teach the requisite content to a diverse class of learners, especially students who have considerable academic ground to make up.

While most people would agree on the importance of all teachers being both qualified and effective, opinion diverges on the value of making this an NCLB requirement. If it is required, how should teacher effectiveness be determined? Would such a requirement be productive or create unintended consequences, such as teachers fleeing the schools that pose the most
difficult challenges? Policymakers need to give careful consideration to the federal actions that would be most constructive in helping all schools get the teachers they need. An important step would be to give attention to what high-quality teaching looks like in various kinds of schools, how to assess it, and how to prepare new candidates and support existing teachers to reach high standards of effectiveness.

In the nation’s need for highly qualified effective teachers, there are both capacity and distribution problems. Despite some schools’ improvement in outcomes for low-income and minority students, these students on the whole continue to have dramatically less access to effective teaching. High-poverty and high-minority schools tend to have more novice teachers, and many of them move on as soon as they get some experience under their belts. This revolving-door phenomenon is especially common among those schools with principals lacking the leadership skills to give teachers the supports and guidance needed to meet the difficult challenges they face.

Evidence suggests that teachers are more likely to stick with a struggling urban or rural school in their own community, and they are often more effective in teaching those students than are “imported” teachers. Based on this finding, some localities are trying the “grow your own” approach to staffing challenges. Such programs identify prospective teachers within the local area—through high school counselors, colleges of education, and self-referrals—and offer supports such as seminars, forgivable loans, work experiences, mentor support, and assistance with job placement. Evaluation of such programs, along with funding if they prove successful, could aid hard-to-staff schools in achieving stable, motivated teaching forces.

Another strategy that merits attention is design of inservice professional development based on best practices shown to help teachers improve low-performing students’ skills. Further, schools might see benefits from giving teachers time during the school day to develop lessons collaboratively and share ways to enhance these based on classroom experience, as in the Japanese model.

**Shortage of mathematics and science teachers**

In many schools and districts, especially in inner cities and rural areas, qualified mathematics and science teachers are in short supply. NCLB now requires schools to show achievement gains in mathematics; a requirement in science may be added next. Without doubt, these are key areas that young people need to succeed in college and in the 21st century workplace. Yet the national teacher shortage in math and science persists. Schools are not producing enough individuals sufficiently grounded in math and science for many of the jobs they encounter in the workplace, and college graduates in these fields are in demand for jobs with higher salaries and status than teaching offers.

An obvious course of action is paying teachers higher salaries in fields where schools have to compete for them in the labor market. Teacher unions have traditionally opposed such differentiated pay, although they may be beginning to accept that differential compensation is appropriate under some circumstances. Besides salaries, other financial incentives to attract teachers to hard-to-staff fields and schools include tax credits and signing bonuses, as well as loan forgiveness, which may be attached to a multi-year commitment. Such approaches are most likely to succeed when schools also provide mentoring and induction programs for teachers new to the school in addition to high-quality professional development for all teachers. Federal funds would allow districts and schools to use such strategies more widely.

To upgrade the level of students’ knowledge of mathematics and science and increase the availability of teachers in these high-demand areas, some policy experts have suggested that Congress consider legislation similar to the National Defense Education Act (NDEA) of
1958. Following the Soviet launch of Sputnik, this bill was enacted to help the nation rapidly make up ground in mathematics and science. Multiplying U.S. investment in university-based research and providing incentives and opportunities to pursue math and science teaching, NDEA helped the U.S. to deal with hurdles similar to those faced today and might serve as a model for new legislation.

**Essential leadership for today’s schools**

Effective principals significantly determine school climate and stability and provide leadership, instructional coherence, and support to teachers. Their leadership makes a major difference in whether teachers choose to come to and remain in a given school. Today good principals are more crucial than ever. Yet the job has become dauntingly difficult to perform. In schools facing the greatest challenges, the principal’s role is particularly tough. Not surprisingly, schools and districts with more intense needs are seeing rapid turnover in school leaders. To meet the aims of NCLB, principals need new skills and greater leadership capacity than ever before.

Unfortunately, most of today’s university-based principal preparation programs are unlikely to give participants the needed skills to meet the challenges of NCLB. The leadership programs that are having the greatest success are quite different from typical university-based programs. One promising model, New Leaders for New Schools, includes a year or more of relevant academic study, intensive skill development in instructional and organizational leadership, and a paid residency in an urban public school. After graduates move into positions as principals, they continue to receive ongoing coaching and mentoring from the program’s staff and alumni, all of whom have made a long-term commitment to urban school leadership. A number of cities have adopted elements of this model for their principal development programs. Federal support for capacity building in school leadership has considerable potential to improve school achievement. The modest expense of incentive bonuses and leadership development for such approaches may be a cost-effective way to leverage student progress.

To be effective, principals also need to be able to budget, hire, and perform other management roles with more room for judgment than the system currently affords them. For example, if principals cannot refuse teacher transfers based on seniority, this constraint may significantly hinder their school improvement efforts. Despite union opposition, some states and localities are taking action to guarantee that principals have this right of refusal. Congress could consider ensuring this right in NCLB to enable principals to drive school progress.

**Accountability and data systems**

NCLB’s insistence on disaggregation of data by groups within the student population (minorities, English language learners, and students with disabilities) has been powerful in pushing states, districts, and schools to raise the bar and consider how to enable all students to succeed. At the same time, certain changes in data and accountability systems may be needed if NCLB’s intent is to be realized. These changes may include use of growth models for measuring progress to proficiency; multiple measures of achievement; calculating and reporting graduation and dropout rates more adequately; and finally, timely release of data.

**Measuring growth to proficiency.** The framework NCLB sets out for annually measuring whether schools are in compliance with requirements regarding Adequate Yearly Progress (AYP) has important implications for schools and the academic achievement of students, especially traditionally low-performing students. Changing how we measure school proficiency, some experts argue, could improve outcomes for all students. It could also help to address the widely held criticism that AYP requirements are unfair.

Under current AYP rules, schools and districts are graded by comparing the scores of grade cohorts of children with an annual target
rather than by calculating individuals’ progress. This method is called a status model. An alternative approach, the growth model, is being piloted in a few states. Growth models measure progress by tracking the achievement scores of the same students from one year to the next to determine whether or not they make adequate progress. Incorporating the growth model in AYP rules would give schools credit for students who are on track to becoming proficient within a reasonable time, such as three years, based on the growth trajectory of their scores.

Having their respective advantages in gauging progress toward NCLB’s goals, growth and status models could be used in combination. Growth models also have value beyond their use for accountability purposes. Measuring changes in individual student achievement over time can provide insights to educators about where and how to allocate resources to more effectively boost student achievement.

**Multiple measures.** Many educators and policymakers contend that basing accountability only on large-scale state assessments in two subject areas is not an accurate reflection of a school’s overall performance. The narrowing of the curriculum to what is tested is also a risk. Basing accountability systems on multiple measures may be a constructive direction to take. These measures might include local assessments, teacher-designed classroom assessments collected over time, portfolios and other indicators of student learning, graduation/dropout rates, in-grade retention, percent of students taking honors/advanced classes and Advanced Placement exams, and college enrollment rates.

One rationale for using a broader array of indicators is that effective school reform must address the realities of work and civic life in the 21st century and thus needs to go beyond the AYP in literacy and math. Today’s workers need to be capable in high-level thinking, problem solving, collaboration, and teamwork. Corporate America sees a focus only on basic skills as threatening our education system and economic viability. Meaningfully assessing 21st century skills will require assessments that measure higher-order thinking and problem solving, utilizing more than multiple choice questions.

**Graduation and dropout rates.** To discourage schools from raising their achievement scores by pushing out lower-performing students, NCLB requires secondary schools to also report their graduation rates as part of the AYP. At present, however, they only need to report a total rate for all students. They do not have to show improvement, or even a plan for improvement, towards a stated goal. High schools could be required to set goals for improving graduation rates and to disaggregate the rates they report for all subgroups of students. Such requirements would clarify whether schools are closing the gap between subgroups on the important indicator of graduation rates.

However, this requirement alone would not bring about desired results. Also key is how states and districts calculate their graduation and dropout rates. Recent studies reveal that self-reported dropout rates and graduation rates by states and districts are highly inaccurate because of variations in calculation methodologies and inadequate reporting mechanisms. Until recently, both federal and state data indicated that high school dropout rates were no higher than 12 percent and that rates for minority and white students had converged. New analyses show that the overall high school graduation rate is much lower than previously assumed and reveal substantial gaps between African American and Hispanic students and their white counterparts. It is now clear that school districts have been reporting rates based on information and methods that were quite inadequate if not deliberately misleading. Fortunately, a rigorous methodology that tracks students annually from the 9th through 12th grades is available, and it yields a far more accurate picture of the dropout rate. The reauthorized NCLB could require every state to adopt an assessment system along these lines and create the data systems needed to track every child.

**Timely release of data.** Despite all the effort and expense that go into student assessment for NCLB purposes, the data are rarely useful to
schools, states, and districts, or to families, because they are not released for many months, often well into the next school year. Requirements for speedier release of data could be combined with the financial support that states may need to upgrade their data systems.

Accountability and children with disabilities

Until NCLB many students with disabilities were excluded from large-scale assessments such as those mandated by states. While it may be too early to tell what impact NCLB will have on these students’ proficiency when fully implemented, it has already changed how people think of issues related to accountability and students with disabilities. The inclusion of all students in the calculation of AYP has forced educators to concern themselves with the education of students with disabilities. The Individuals With Disabilities Education Act (IDEA) was a groundbreaking legislative instrument that supported the teaching and learning of students with disabilities. It was not until NCLB provided the accountability mechanism, however, that educators began to see the capabilities of these students on a wider scale.

Like all other students in the schools, those with disabilities need access to the general curriculum, qualified teachers, and high expectations for their learning. When they receive appropriate accommodations, services, and supports, an estimated 85% to 90% of these students can be expected to achieve grade level mastery on the achievement tests currently in use. An adjustment to NCLB permitted states to develop alternate assessments of proficiency and administer them to a maximum of 1% of the entire student population (about 9% of the disabilities subgroup), those with significant cognitive disabilities. To date, there is no evidence to suggest that it would be advisable to raise this level from 1% to a higher proportion of students. Further, experiences with the students in the 1% group have shown that there are dramatic benefits to them through the requirement that they be assessed and that their performance be reported separately. Substantial research having developed around this type of assessment, systematic methods now exist to ensure them greater access to the general curriculum.

There is some concern that disaggregating data and holding schools and districts accountable for the performance of subgroups has created scapegoats, blame, or negative attitudes toward low-performing populations. This issue is not to be discounted. Yet NCLB has done much to bring attention to students with disabilities, their need for access to the curriculum, for individualized services, supports, specialized instruction, better assessments, and appropriate accommodations in testing. While it may be challenging for schools to rethink their practices for students with disabilities, they are beginning to do so.

Coherent standards across states

Under NCLB each state sets the bar for what constitutes proficiency on its own tests. Examining results on the common metric provided by the National Assessment of Educational Progress (NAEP) reveals wide state-to-state differences. Further, NAEP results show alarmingly high proportions of students performing below the proficient level in reading and other key areas, far more than the proportion scoring below the proficient level on state tests used for NCLB accountability. Most state assessments clearly set the bar too low. In fact, no state has a system of standards, curriculum, graduation requirements, and accountability that is well aligned with the demands of postsecondary education and work.

The United States has a decentralized education system, but nonetheless needs to achieve consistency in standards nationwide. The knowledge and skills young people need today are not markedly different across state boundaries. Also essential is for all standards to clearly define what students should know and be able to do at different grade levels. In turn, these standards should be based on the realities of
what students eventually need to know and do in postsecondary education and the workplace.

This need for coherence and consistency does not necessarily mean we need a national curriculum or top-down national standards. Another route is being pursued by the American Diploma Project initiated by Achieve, Inc. Achieve is a bipartisan organization created in 1996 by the nation’s governors and business to help states raise academic standards, improve assessments, and strengthen accountability in order to prepare all students for postsecondary education, work and citizenship. The American Diploma Project Network is a coalition currently made up of 29 states committed to aligning K-12 curriculum, standards, assessments and accountability policies with the demands of college and work. At present, high schools are not being held accountable for students attaining college and workplace readiness. K-12 and higher education should be partners in ensuring that all young people are prepared for work and learning. Both have critical roles to play in ensuring that students who go on to postsecondary education stay in school and graduate.

Funding for NCLB

No Child Left Behind is often referred to as an unfunded or underfunded mandate. Although most educators and policymakers would agree that far more effective use could be made of existing resources, there is also considerable consensus that the public schools cannot meet the highly ambitious goals of NCLB without higher levels of federal funding. Opinions may differ as to whether “full funding” of NCLB is essential, but at a minimum, the states will need to see the federal government is a reliable partner. At present, many people see the federal government as having handed states and districts the costs of meeting an ambitious mandate rather than providing the supports and resources necessary to accomplish its goals. The issue of increased funding will be debated at length as reauthorization goes forward.

It is also important to consider the funds that are now in the system but are not getting where they are intended to go, notably Title I funds. These dollars are supposed to target high-poverty, high-need schools, yet many end up in suburban schools because of what is called the “comparability loophole.” Many districts, in effect, hide the school-to-school inequity in teacher resources by averaging teacher salaries across all their schools as if every teacher were paid the same. This practice could be prevented by requiring that salaries be reported by school rather than averaged across the district to ensure that Title I and non-Title I schools have similar expenditures for teacher salaries. Congress should consider any potential measures to ensure effective targeting of resources to high-need schools.
My task is to discuss the relation between the U.S. educational system and its economic competitiveness, a topic on which far more needs to be said than I can cover in these brief remarks. As I shall argue, the long-term link between the performance of the education system and of the economy is strong. Nonetheless, I begin with three caveats.

1. Americans are rightly concerned with the gap in academic achievement between whites and minority groups such as African-Americans and Hispanics. Closing that gap is the principal goal of the No Child Left Behind Act, signed into law by President Bush in early 2002 and up for reauthorization this year. But it is important to acknowledge that fully half of the gap we see in 12th grade is attributable to early observable differences when students enter 1st grade. Differences in household income, parental education, and family structure have a significant impact on children’s readiness to learn. And regrettably, the link between family background and student performance is stronger in the United States than in most other advanced nations.

2. We have an educational problem that takes precedence over the achievement gap—namely, a dropout crisis. While there is a complex debate now underway among scholars, it is already clear that actual dropout rates are significantly higher than federal statistics have indicated, and that there are wide gaps among white and minority students. This matters enormously. Labor force participation among dropouts is only 42 percent, versus 60 percent for high school graduates. Those dropouts who are employed on average earn 30 percent less than graduates. The annual economic impact—measured in earned income and tax receipts—amounts to nearly $200 billion (about 1.6 percent of gross domestic product). If one adds in the social costs—much greater rates of incarceration, substance abuse, and homelessness, among others—the social costs are higher still. And to complicate matters further, it turns out that students who complete their high school degrees via alternative paths—either returning to school after leaving or fulfilling the requirements for the GED—have incomes closer to those of dropouts than of those students who stayed in school and got their diplomas on time.

3. Education is but one of many factors influencing international competitiveness. Others include macroeconomic policy, the size and direction of public investment, legal and regulatory frameworks, and levels of research and development spending in the private sector.
Let me now return to my assigned task. I begin with the impact of education on the relative position of individuals within the U.S. economy and then explore its impact on the relative position of nations.

**Individuals.** Over the past 30 years, the premium that college graduates in the United States enjoy over high school graduates has more than doubled (from 40 to 80 percent), and that gap is much higher here than in any other advanced industrialized nation. During this period, the real incomes of college graduates rose by 19 percent (holders of advanced degrees enjoyed a 46 percent increase) while high school graduates lost about 1 percent and dropouts fell a full 15 percent. Not surprisingly, this enlarged gap has had a major impact on economic mobility over the past three decades. Individuals with a high school degree or less have tended to fall out of the middle class, while those with some college or more have tended to enter or remain in it, or even to attain upper middle class status. It turns out, moreover, that job quality multiplies the income effect of educational differences: holding education constant, jobs requiring higher levels of critical thinking, innovation, and advanced problem-solving pay more, and that trend appears to be accelerating.

**Nations.** For much of the 20th century, the United States had the best-educated workforce in the world. That is no longer the case. Many nations now exceed the United States in high school graduation rates, and during the past 30 years the United States’ share of the world’s college graduates has fallen by more than half. Moreover, the pace of our educational gains is slowing dramatically. In the two decades from 1980 to the end of the century, the percentage of the U.S. workforce with at least some college rose by 20 points. Given current trends, the gain from 2000 through 2020 will be at best 4 points. Meanwhile, the entrance of China, India, and the successor states of the former Soviet Union into world markets has increased the global labor supply by about 2.7 billion. Economic competition has intensified and will continue to do so.

International comparisons of educational performance place the United States in the middle of the pack ... at best. The Trends in International Math and Science Study (TIMSS) administered in 2003 showed that U.S. students were 6th among 11 OECD countries in 4th grade math and tied for 6th in 8th grade math. U.S. Science attainment was somewhat better: tied for 2nd at the 4th grade level and tied for 4th at the 8th grade level. The Program for International Student Assessment (PISA) revealed a gloomier picture: U.S. students were tied for 11th (out of 29 nations) in science, 13th in reading, 21st in math, and a dismal 23rd in problem-solving.

Sophisticated econometric studies indicate that both the quantity and quality of a nation’s education affect its overall economic growth rates. A one grade increase in average years of education translates into roughly a 0.5% increase in real GDP/worker. In the United States, a 10 percent increase in the college completion rate (from 25 to 27.5%) would produce an estimated one percent increase in real GDP/worker (roughly $125 billion in 2007 dollars). While the scholarly debate over the relative impact of educational quantity and quality is ongoing, the eminent economist Robert J. Barro reaches a conclusion that seems to me consistent with the bulk of the reliable evidence: “the quality and quantity of schooling both matter for growth but ... quality is much more important.” This suggests that as the U.S. edge in quantity of education erodes or disappears altogether, the relative quality of our educational system will matter more and more. And given our mediocre educational performance relative to our peers, it is hard to escape the conclusion that our more than respectable economic growth rate is attributable for the most part to features of our policies and practices other than education.

Summing up: the United States could grow faster, and with less inequality, if our education system worked better—if we had higher achievement in K-12 public education with fewer high school dropouts and small gaps
between groups and if we had higher rates of college attendance and completion. Whether we can achieve these goals with a system as fragmented as ours at present strikes me as a matter worthy of the most serious discussion by our elected officials. Recent signs are not encouraging. In February 2007, the National Assessment of Educational Progress (NAEP) released a report showing that the reading skills of 12th graders tested in 2005 were significantly worse than those of students in 1992: the share of 12th graders lacking even basic reading skills jumped from 20 to 27 percent, while those rated as proficient fell from 40 to 35 percent. In math, 39 percent of 12th grade students lacked even basic skills. Students registered these dismal results despite averaging 360 more hours of classroom instruction in high school than had their peers in the early 1990s.

What should we do? I believe that policymakers should give special attention to:

• universal pre-K and other programs to ensure that many more students reach public school ready to learn;

• programs that identify, no later than the eighth grade, students in danger of dropping out and that designate mentors to guide these students through their high school years;

• financial assistance packages that replace loans with grants for low-income college students; and

• a plan for eliminating the gap between the rigorous NAEP exams and the much weaker state standards permitted under the 2002 No Child Left Behind Act.

One way or another, we must come to grips with two basic truths. First, competence in core academic subjects is the same, wherever you live. Math is the same in Chicago and Cheyenne … and for that matter, Shanghai. And second, in a global economy, competence means measuring up to standards of excellence established, not by 50 chief state school officers, but in the capitals of our competitors around the world. While we may continue to reject a more centralized system of public education as antithetical to our traditions and values, in the end we will have little choice but to move toward much more uniform and rigorous standards of academic competence for our public schools. The only remaining question is how long it will take our citizens and their elected representatives to accept these truths.
**Background**

The Cleveland Municipal School District is comprised of 120 schools that serve 69,532 students. The student body is 70.6% African American, 17.7% Caucasian and 9.5% Latino. One hundred percent of the students are economically disadvantaged, according to the USDA standard qualifications for free and reduced lunch. Cleveland has been declared the poorest urban city in the country for two consecutive years by the United States Census Bureau. Like other urban districts, the Cleveland Municipal School District faces difficult challenges. The state school funding system has been declared unconstitutional by the Ohio Supreme Court four times, but has remained intact, resulting in inequity and financial difficulties.

In 1995, U.S. District Judge Robert Krupansky disbanded the Cleveland School Board and instructed the State Superintendent of Public Instruction to assume immediate fiscal and operational management of the district. Judge Krupansky took extraordinary action because by every conceivable measure—educational, financial and organizational—the Cleveland Public School District was bankrupt.

Only one third of the students enrolled in the eighth grade graduated. Test scores across all grade levels were among the lowest in the state; just 15% of 4th graders and 11% of 9th graders passed the Ohio State Proficiency Tests. On any given day 20% of the students were absent from class and the District’s financial situation was out of control. The school district had accumulated debt of $152 million and the cost of needed capital improvements was estimated at nearly $600 million. The average tenure for the superintendent was 20 months.

In 1998, then Cleveland Mayor Michael R. White and the State of Ohio worked with civic, business and neighborhood leaders to focus on the Cleveland school system. By September of the same year, the Ohio Legislature passed HB 269, mandating a new governance structure for Cleveland schools. As a result, Mayor White gained control of the school system. In September, 1998, he appointed the new school

---

No Child Left Behind: Revisiting Adequate Yearly Progress

Barbara Byrd-Bennett, PhD
Executive-In-Residence
Cleveland State University

Our nation has a moral imperative to close the achievement gap between low income students and their more advantaged peers. The No Child Left Behind Act makes this a legal requirement as well. Yet improving learning opportunities for all children will require more than individual talents or school-by-school efforts. It will demand system-wide approaches that touch every child in every school in every district across the nation.

(Beyond Islands of Excellence: What Districts Can Do to Improve Instruction and Achievement in All Schools, Togneri and Anderson, 2003, p. 1)
board and after an extensive search Barbara Byrd-Bennett was installed as the Chief Executive Officer of the Cleveland Municipal School District.

Prior to serving as the Chief Executive Officer of the Cleveland Municipal School District, Barbara Byrd-Bennett was appointed by then Chancellor Rudy Crew as the Supervising Superintendent of the Chancellor’s District in New York City. The Chancellor’s District is composed of the lowest performing schools in New York City.

A New Beginning

The introduction is meant to provide a live context for the No Child Left Behind Act. Although the Act is ambitious, most urban superintendents generally agree with the fundamental principles. For the past two decades, the public has expressed its concern with both the real and perceived failure of public education in America. As a result of the public’s growing frustration, an environment of consensus evolved which allowed for a bipartisan Congress to pass the No Child Left Behind Act in 2001.

The Act represents an unprecedented expansion of the role of the Federal government in public school education, kindergarten through grade twelve. The core provisions of the No Child Left Behind Act are embedded in the principles of raising student achievement, closing the achievement gaps between and among different groups of children, improving the qualifications for all teachers in all classrooms, developing state standards and state annual standardized testing systems for accountability, providing students with tutoring and supplemental education services and creating opportunities for parental choice.

There is little disagreement or debate among educators about the necessity, timeliness and appropriateness of the Act and most of the provisions. Most superintendents would agree that there is a need to require states to set academic standards for all students and to further require states to develop a system of annual testing to measure the achievement of all students at each grade level. This provision helps to erode the historic complacency about the academic performance of poor, African American and Latino children attending public schools throughout this country.

Not since the Brown v. Board of Education Decision (1954) has there been focus on the question of equity of access to learning opportunities for all children enrolled in America’s public schools. More than fifty years after the Board decision, the No Child Left Behind Act is a welcomed catalyst for thoughtful discussion and immediate action to improve the quality of teaching and learning and to accelerate academic achievement for previously ignored groups of children.

The requirement for states to set academic standards for proficiency, to develop a standardized testing system and to annually report the results of those tests to the public, make public education more public. As a result, the necessary transparency exists for everyone to compare the performance of different groups of children to each other within school and to compare the progress of groups of students in schools throughout the school district. It helps everyone to understand the need for redirection of existing resources and/or the need for additional directed resources. This is a very general statement of approval because we have learned that compliance with particular provisions can create false perceptions about public status of the progress of public schools. In addition, each state creates its own regulations for adherence to a particular provision.

The No Child Left Behind Act does expose some of the myths that have survived for decades. Low expectations and standards coupled with poor teaching result in low achievement. No longer can issues of race, income, residency and family background be accepted as excuses for poor results. No longer can children who were previously made comfortably invisible remain invisible.

This year, school districts and schools across
the country are beginning to feel the effects of
the federal legislation as we move closer to the
deadlines imposed by the law. Now is the time for
reflection and assessment of the core provisions.

Like most urban superintendents, I believe
that the No Child Left Behind Act has been pos-
itive for public education, but I do not believe
that the Act will directly transform public edu-
cation and solve all of the social and economic
problems that impact education.

Like most of my colleagues, I do not fear
accountability. I understand accountability and
I have always been and will continue to be held
accountable for my work.

Like most of my colleagues, I am concerned
with the reality of reaching the goals of the leg-
islation given the current circumstances, issues,
obstacles and challenges which confront
urban school districts.

Given the schools and school districts where
I have worked, I am profoundly aware of the
challenges of poverty, family mobility, substan-
dard housing and inadequate health care and
the implications of each on our ability to edu-
cate students. I am also painfully aware of the
impact of limited financial and human capaci-
ties of urban school districts. Although these
limitations do not make the goal of the No
Child Left Behind Act less important and less
urgent, the limitations are real. They do make
reaching the goals of the No Child Left Behind
Act more difficult and seemingly more illusive.

Nevertheless, superintendents, principals
and teachers are struggling to meet all of the
complex details and stringent requirements of
the law within the prescribed timelines.

Flawed Framework for Adequate
Yearly Progress

Despite enormous expense and effort, no
statewide systemic improvement plan to dra-
matically accelerate student achievement in
public schools has been developed and imple-
mented. School districts across this country are
very similar to governments. Each struggles
with issues that are layered and complex and, as
such, the causes of the problems are not easily
remedied by a “one size fits all” solution.

However, the key accountability provision
mandated by the No Child Left Behind Act is
Adequate Yearly Progress. It is a “one size fits
all” provision requiring states to develop a uni-
form accountability system for all districts and
schools within its jurisdiction.

In its current form, the framework for
Adequate Yearly Progress requires children to
achieve mastery on state reading and mathemat-
ics tests in grades three through eight. This
school year, it also requires state examinations in
science be given at the elementary school level.

In addition to the required tests under the
provision, there is also a requirement that
“other academic” indicators be used. At the ele-
mentary school level, attendance rates serve as
an “other academic indicator.”

Students at the high school level are required
to successfully complete selected basic skills
tests. In most instances, these tests are required
to be administered by tenth grade. At the high
school level, the graduation rate serves as an
“other academic indicator.”

The Adequate Yearly Progress provision is
based upon the assumption that testing and stan-
dards will improve low performance. It further
assumes that the causes for low performance are
the same for each state and school district. The
cause of the problem has been incorrectly diag-
nosed, so the solution framework is flawed. As
such, progress will be marginal.

Schools that were previously deemed effec-
tive and schools that have been identified by
the U.S. Department of Education as Blue
Ribbon Schools now find themselves on the fail-
ing schools list in their states. The number of
school districts and schools that have failed to
meet Adequate Yearly Progress has increased
since 2003 and is expected to continue to grow.

Although monitoring student progress on
standardized tests and adult accountability for
results is an essential provision to the federal
law, as it is currently structured the Adequate
Yearly Progress provision oversimplifies the
complexities of the problems faced by school districts and schools. Under the provision, accountability is determined and measured by student test results. Student academic progress is measured by comparing the test scores of one grade in a particular subject to the scores of the same grade in the same subject from the previous year. The fourth grade reading test results for 2005 are compared to the fourth grade reading test results for 2004 to determine improvement and progress. Individual student growth and progress is not a factor in the formula and as a result too many children are being left behind.

The same method of comparisons is used for “subpopulations” of students. The “subpopulations” of students are represented by different groups of students. Some of the “subgroups” include students with disabilities, poor students, African American, and English Language Learners. Every student, including those categorized in the “subgroup,” is expected to reach mastery on the state standardized tests by 2014.

Another complication of the flawed framework is the “safe harbor” provision. Schools that make progress, but fail to meet the targets established by the state, can meet a “safe harbor” provision. This is accomplished if at least ten percent of the students annually score at a mastery level on the state’s proficiency test.

Districts and schools are accountable for the year-to-year results of standardized test scores. If they fail to meet the expected targets each year, the district and/or school will be identified as ineffective or failing to meet Adequate Yearly Progress. If a district or school is deemed as failing, it will be subject to imposed sanctions. The sanctions begin with restructuring and move through state takeover, charter school conversion, management by a private company and withholding of Title I funds.

Based upon the current Adequate Yearly Progress provision, the districts and schools that enroll the largest number of the neediest students with the greatest diversity suffer the harshest punitive sanctions. The heavy-handed approach to schools that fail to meet Adequate Yearly Progress is arcane and counterproductive to the goal of the No Child Left Behind Act.

Challenges for the Field

Most superintendents have one goal and that is to improve student achievement. The struggle is to ensure that rigor exist and that the standards are high and aligned to the assessment. Student performance can only be enhanced and sustained by quality teaching and not certification credentials. And ultimately, we need to restore a belief in a system of public education for all children.

The question then is: How far does the school need to advance different children to be deemed effective? What are the implications for using absolute standards versus relative standards? We know that in the lowest low performing school districts, grade-to-grade, year-to-year comparisons do not explain student progress. Other significant factors such as student mobility, language diversity, teacher quality, changes in grade configuration and composition of the classes each year at a given school impact the data.

The method for collection of the achievement data must be changed. In order to create a meaningful system for accountability, a data collection system capable of tracking and benchmarking student achievement is needed. We know the further behind a student is from the target, the greater the resources and interventions will be needed. But, what those resources are and how they are working can only be determined by a system that evaluates and assesses the inputs.

A testing system must be aligned to rigorous and high standards and must also be diagnostic. Relative student achievement or a value added model is a more valuable measure of school effectiveness or ineffectiveness. A system that measures student growth and progress over time is real accountability.
Given the changing demographics in the United States, any education policy must be thoughtful about the notion of “subgroup” populations. Nearly 43% of the students enrolled in our public schools are minority and/or their first language is not English. In large urban school districts, minority students are overrepresented in special education classes. Districts need greater flexibility for determining test exempt students. This would include, under certain circumstances, special education students and English language learners.

There is wide variation among states for group size. The group size effect can unjustly and negatively impact a school’s rating.

No one group score should cause an entire school district or school to fail.

If there is one lesson that I have learned after more than thirty-five years in public education, it is that our changing world is at once unknown and unknowable. The challenges that face the next generation of schools will require the continued spirit and principles exemplified in the No Child Left Behind Act. Schools and their communities are inextricably bound and the health of the community is determined by the health of its schools. The challenges of education that our country will face in the future can eventually be solved with careful, coordinated discussions and policies that reflect a recognition of the complications of diversity and the perils of the poverty found in many of our inner cities. The future greatness of this nation rests solidly in the present quality and the future potential of our schools.
Measuring Proficiency With Status and Growth Models: How Best to Maximize the Achievement of All Students?

Jim Wyckoff, Ph.D. *
Professor of Public Administration, Public Policy and Economics
University at Albany

No Child Left Behind sets out a framework for annually measuring whether schools are in compliance with the Act’s requirements regarding Adequate Yearly Progress. This framework has important implications for schools and the academic achievement of students, especially traditionally low-performing students. The past five years have provided valuable experience with the mechanics of NCLB. This paper explores whether changes in how school proficiency is measured could improve outcomes for all students.

The No Child Left Behind Act (NCLB) seeks to insure that all K-12 students meet high learning standards and that the achievement gaps between student subgroups are eliminated. NCLB provides a framework to measure progress toward these goals and an accountability system for schools with substantial sanctions if the goals are not met. The definition of proficiency and how states measure the progress of their schools toward proficiency have important consequences for states’ compliance with NCLB and, more importantly, for the growth in student achievement. Although the measurement of proficiency may appear to be a technical detail, it has important implications for the success of NCLB.

This paper explores the implications of employing the current standard of measuring school progress toward universal proficiency, the so-called status model, with an alternative approach currently being piloted in a few states usually referred to as a growth model. In short, growth and status models measure different aspects of changes in student achievement, both of which are valuable in gauging progress toward NCLB’s goals. However, within the context of NCLB both status and growth models suffer from potential problems. Growth models are useful beyond their value for accountability. Measuring changes in individual student achievement over time can provide insights to educators about where and how to allocate resources to more effectively improve student achievement.

The paper begins with an example of student knowledge and how it can change over time that illustrates some of the differences between status and growth models. The next section describes how status models measure proficiency for compliance with Adequate Yearly Progress (AYP) and some of the issues that have been raised about the use of status models. This is followed by a description of how growth models measure proficiency and how they could address the concerns raised regarding status models. Finally, the paper explores some of the practical issues associated with employing growth models.

Gains in student knowledge.
Let’s consider a typical example of the achievement of students in a school and how achievement changes over time. Suppose that the actual
The math knowledge of three specific students is also shown. Connie (C) knows substantially more than both Bob (B) and Alice (A) but Alice is more typical of students in the 4th grade at PS10. Suppose 4th grade proficiency is shown by the vertical line at P4. Only 40 percent of the students are proficient as most of the area under the curve is to the left of this proficiency level. In this instance neither Bob nor Alice is proficient.

One way of examining how knowledge has changed at PS10 between 2005 and 2006 is to examine how a new group of 4th graders performs. Fourth graders in 2006 are shown by the lighter curve in Figure 2. The rightward shift in the curve reflects an improvement in knowledge, or status. Now, 50 percent of the fourth graders are proficient. However, because this is a different cohort of students, it is difficult to assess whether the school is responsible for the improvement or whether this is just a more able group of students.
A different approach to examining the change in knowledge is to follow the 2005 cohort of 4th graders over time to assess increased achievement. As shown in Figure 3, these students know much more math in 2006 than they had in 4th grade as evidenced by the movement to the right of the dashed curve. Each of our illustrative students has improved, some more than others, as indicated by the length of the arrows. Alice experienced a much larger gain than Connie and is now proficient. Bob had a large gain but fell short of 5th grade proficiency. Note also that by 5th grade overall student proficiency has improved to 50 percent.

This example illustrates two important aspects of student knowledge. First, status and growth are very different concepts and provide different information. For simplicity imagine that achievement in a school can be represented by low or high status and low or high growth. All four combinations (high status, high growth; high status, low growth; low status, high growth; and low status, low growth) are found in empirical observation of schools. So, a school could have mostly high achieving students (high status) who make relatively small gains in knowledge (low growth). Another school may experience the reverse. Second, a proficiency benchmark measures progress regarding one specific level of knowledge. As a result, it is much easier for some students to reach proficiency than others (Connie v. Bob). This section describes different ways of describing changes in knowledge over time. How changes in achievement are summarized can have important implications for accurately describing student knowledge and for realizing the goals of NCLB.

Measuring Student Knowledge with Status Models.

NCLB sets out the metrics by which states gauge the progress of schools toward proficiency for all students by 2014. All states must use a status model to measure the knowledge for all students and subgroups of students in each school. The key components of a status model are:

- States determine the learning objectives and level of knowledge in math and reading/English that constitutes proficiency in grades 3 through 8 with additional subjects and grades included over time.
- States set annual measurable objectives (AMO) that indicate what portion of the students must be proficient in each year until 2014 when all students must be proficient. States have flexibility in how they determine the AMO, but a typical example determines the percent of students statewide who were proficient in 2002, calculates the difference between this proficiency level and 100 percent and then
divides the necessary progress evenly over the intervening 12 years. For example, suppose 40 percent of the students in State A were proficient in 2002, implying that the remaining 60 percent need to become proficient by 2014. Reflecting the target of an additional 5 percent becoming proficient each year, the AMO for 2003 would be 45 percent, the AMO for 2004 is 50 percent, etc.

- States measure the proportion of all students in each school who are proficient based on annual student achievement tests. If the school-wide and student subgroup (i.e. economically disadvantaged students, major racial/ethnic groups, students with learning disabilities and students who are Limited English Proficient) proficiency rates are at least as great as the AMO for that year, the school has met this component of AYP. Groups with sample sizes below a minimum level are excluded from AYP.

Concerns have been raised about whether status models provide good information regarding schools’ compliance with NCLB goals. These concerns include:

a) Misleading indicators. A status model may provide misleading indicators of a school’s improvement.

- Since the composition of a school’s students changes over time, a school may increase its proficiency rate, even though the school may not have improved its ability to educate children.
- A status model does not account for the extent of achievement gains students must make to reach proficiency. Thus, all the students in a school may make large achievement gains, but the school may not reach the AMO because its students began so far from proficiency (Bob in Figure 3 is an example).

b) Incentives. NCLB provides strong sanctions if schools do not reach proficiency goals. The high stakes may create unintended incentives for schools.

- In the short run, schools have an incentive to emphasize the progress of students just below the proficiency threshold (the so-called bubble kids), such as Alice, placing less emphasis on the worst performing students thereby leaving the lowest performing students behind.
- Schools also have incentives to structure their composition to enhance their ability to reach the AMOs, such as limiting incoming transfers to higher performing students and reclassifying students among subgroups like special education and Limited English Proficient.

- Once a student reaches proficiency the status model usually provides little incentive to improve achievement beyond this threshold. As a result, good students, e.g., Connie, may be neglected, resulting in under-achievement or student migration to other schools.

c) Instructional Assistance. Status models indicate whether students are proficient or not, but do not identify which students are having the most difficulty in making achievement gains. Thus they provide only limited insight on how to target resources to improve the teaching and learning environment.

d) Unfair accounting. The status model may be an inequitable accounting of school performance as it confounds school and student factors. Schools with students whose family background and prior education result in high initial achievement may not need to make much effort to reach proficiency goals; other schools whose students are not as ready to learn can make enormous progress and still fall short.

e) Group size and test error. As the size of schools and student subgroups gets small, it becomes more difficult to meet AYP. A bad
performance by just a couple of students substantially affects the proficiency rate. This issue is compounded by error in measuring student performance. Tests are subject to error for a number of reasons, e.g., a limited sample of questions and idiosyncratic events at the school such as the flu or a faulty heating system. Therefore, the results of any test administration may misclassify students with respect to proficiency. For purposes of NCLB, many of the individual errors will net out as students are aggregated to schools, but this is less likely the smaller the group size.

However, there is substantial variation in the group size levels adopted by the states. Some states have minimum group size of five or less, while others have group sizes of fifty or more. This issue relates closely to that of the use of confidence intervals in determining whether schools or subgroups have met the annual measurement objective. Here, too, there is wide variation in how states have employed this discretion. Taken together, variations in the implementation of group size and confidence intervals have led to substantial differences in the measurement of student achievement across states.

f) **Teaching to the test.** Assessments focus on some knowledge in math and reading/English which, while important, is a limited subset of what most people believe students should learn. Unfortunately, attaching high stakes to tested outcomes appears to reduce attention to other outcomes. Thus even schools making good progress toward proficiency in some areas of math and reading/English may actually be reducing proficiency in other, non-tested areas and in other subjects.

g) **Sanctions.** If a school or any of its subgroups has proficiency rates below the AMO, that school is subject to sanctions which increase if the school remains out of compliance in successive years. This is true whether only one of its subgroups narrowly missed the AMO or if the school and all of its subgroups had proficiency rates substantially below the AMO.

h) **Universal proficiency.** Some have called the NCLB goal of universal proficiency to high standards by 2014 unrealistic and counterproductive. This goal may induce states to set proficiency levels low in order to ease their ability to comply.

### Using Growth Models.

In November 2005 the U.S. Department of Education created a pilot program that could authorize up to ten states to measure progress toward proficiency using models of student growth. Tennessee and North Carolina received approval in May 2006. In November 2006 Delaware was approved and Florida and Arkansas received conditional approval.

The Department of Education requires that growth models employed for NCLB satisfy certain criteria, but has allowed states leeway in their design to facilitate experimentation. Notable among the criteria is that growth models not include controls for student background, thus insuring that all students are held to the same standards. States employ the same definitions of proficiency and AMO as in the status model. The difference is in how achievement test results are used to determine whether a student is proficient. Unlike a status model that determines proficiency based on current achievement levels of students in the school, a growth model allows states to use measures of achievement growth of students over time to determine whether students are on track to be proficient if they experienced that growth annually over three or four years. For example, in North Carolina a student is counted as proficient this year as long as they gain at least one-quarter of the shortfall between their starting level and the proficient score. Suppose proficiency is 300 and a student’s initial score is 240. As long as that student scored 255 in the second year, 270 in the third, and so on, they would be
counted as proficient. For example, although Bob was not proficient in a status model, he could well be counted as proficient if a growth model were employed (Figure 3).

Another type of growth model, often referred to as a value-added model, uses student-level data, but specifically includes the attributes of students such as their background and prior achievement, and statistical controls to separate student and school factors that contribute to achievement growth. When properly estimated, the results of value-added models allow policymakers to distinguish among the many factors that may influence student achievement. Thus, although value-added models are not allowed under current policy they nonetheless may be very useful to state and local policymakers.

Measuring individual student growth can address several of the concerns leveled against status models, although some of the concerns result from generic issues of measuring student outcomes.

a) Misleading indicators. The gains of each student can be compared to those required for proficiency. Since individual students are tracked, growth models avoid the issue of the changing composition of the school. In 2004-05 allowing Tennessee schools to meet AYP by employing either a status or growth model reduced the number of schools that did not make AYP by 13 percent relative to using only the status model. In North Carolina, the reduction was 4 percent.

b) Incentives. Growth models can be designed to avoid the incentives that result from a fixed performance threshold, such as a focus on students near the proficiency threshold, but incentives to select or classify certain types of students remain.

c) Instructional Assistance. Students with low status but high growth are likely to need less intensive intervention than those with low status and low growth. Thus accounting for growth can lead to a more efficient allocation of resources. In addition, although not required for NCLB compliance, value-added models can provide state and local policymakers useful insights regarding school policies that would be most useful in raising student achievement. For example, principals in low-achieving schools could isolate the relative effectiveness of teacher verbal ability, teacher preparation, and teacher mentoring in raising the performance of low-achieving students and thus better target scarce resources.

d) Unfair accounting. Growth models approved by the Department of Education are not allowed to account for student background and thus will continue to confound school and student contributions to achievement growth. Value-added models are able to disentangle the performance of a school from its students. Without sacrificing the goal of universal proficiency, information from value-added models could be employed to distinguish between schools that had strong achievement growth conditional on the readiness of their students from those whose achievement growth was below average. Both can fail to make AYP, but for different reasons. The first may need more resources, while the second may just need to make better use of its resources.

e) Group Size and Test Error, Teaching to the Test, Sanctions and Universal Proficiency. These issues apply irrespective of the model used to measure proficiency.

Data Requirements for Growth Models.

The insights gained from growth models come at a cost. Measuring student-level growth in achievement demands much more of the assessment and the data system than does measuring school-level status. Assessments are judged based on the degree to which there is support for the interpretations of the test scores (validity) and the degree of consistency of assessment results across varying test administrations (reliability). For growth models to
score high on these criteria, it is generally believed that:

- Assessments must be reasonably accurate over the entire range of tested knowledge, not just near proficiency. Additional questions are typically required to accurately measure gains across the expanded range.

- Assessments must be more sophisticated to measure a range of knowledge that spans grade levels and must do so on a consistent scale so that changes at one point represent equal gains in achievement as the same change at another point (vertical alignment and vertical scaling). This often means that there are overlapping questions on tests in adjacent years.

- Databases must link annual student-level test results, socio-demographic information and data on student mobility over time.

- Well executed growth models require skilled statisticians to develop and estimate the models.

- Although not required for accountability, databases that link teachers who taught the class to test results for each student allow the estimation of value-added models used to improve resource allocation.

Creating these databases so they address NCLB compliance and state and local accountability and policy development interests pose challenging substantive, technical, political and economic issues. For example, are standardized tests able to capture the range of knowledge and skills that most people believe students should possess and does testing on a much narrower range create unintended consequences of teaching to the test? Although nine states are employing growth models for accountability and policy purposes, there is debate as to whether the assessments and data systems meet needed high standards of validity and reliability. Some test experts remain concerned with the ability to develop tests that are vertically scaled consistently to permit their use in growth models. Other experts believe the tests employed in the states with growth models approved by U.S. Department of Education meet such criteria. States need to create the technical capacity to uniquely identify students and teachers and link students to the teachers who taught those students. In addition, appropriate concerns regarding confidentiality and privacy need to be addressed. Creating and maintaining longitudinal databases can be expensive. There is only limited information on such costs, but development costs are estimated to be $10 to $20 million per state with additional costs for the annual administration of tests and maintenance of databases.

Summary.

The intent of NCLB is to hold schools accountable for all children reaching proficiency by 2014. This is an ambitious goal with far-reaching consequences. Assume for the moment that this is an obtainable goal when proficiency is interpreted as a high learning standard. How should the federal and state governments measure AYP to create the greatest likelihood that the goal is attained?

Status and growth models each provide important and different information about the achievement and growth in achievement of students. Because status models may provide an inaccurate description of a school’s progress toward universal proficiency, creating the option to also meet AYP using an appropriate growth model is attractive. In addition, growth models provide an ability to tailor accountability to the challenges that each school confronts without sacrificing the ultimate goal of universal proficiency. Growth models also have the potential to usefully inform state, district and school policies intended to improve student outcomes. However, much remains to be learned here and the sites piloting growth models provide valuable experiments about how to structure and use growth models. The U.S. Department of Education should be encouraged to evaluate and disseminate the promising practices.
References


Quality Counts 2007, Education Week, 26(17)


Linn, R. “Validity and Reliability of Student Assessment Results” Working Paper, University of Colorado at Boulder.

* I appreciate useful advice from Dan Fallon and helpful comments on an earlier draft of this paper from Ellen Foster, Hamp Lankford, Susanna Loeb and Kristy Michel. Any errors are solely attributable to the author.
Ten years ago when I took the deanship of the College of Education at the University of Illinois at Chicago, the senior program officer at the MacArthur Foundation challenged me to think about doing the business of teacher preparation differently than usual. He asked me what it would take to really make a difference in preparing the teachers the Chicago Public Schools (CPS) needed. Following are brief retrospective snapshots of the path I and my colleagues have taken in the past 10 years, each snapshot accompanied by a lesson learned.

**Thinking “systemic” and staying focused in a complex and over-regulated environment.**

I knew we had to think systemically and in collaboration with our Arts and Sciences faculty about the entire complex enterprise of quality teacher preparation as it related to providing the teachers Chicago needed most—from initial recruitment and selection of teacher candidates through teacher preparation, including practice teaching in various field sites. We also had to think beyond graduation and certification to beginning teacher induction and mentoring and continued professional development. While I don’t intend to dwell on this, to give a sense of what I mean by “complex and over-regulated environment,” it is worth mentioning that, during this period, we underwent one Illinois State Board of Education (ISBE) accreditation review of our certification programs. We considered whether to strive for either of two national accreditation reviews (we decided against this), and are coming up on a second review this spring, after three lengthy ISBE delays. We underwent an ISBE-mandated conversion of every one of our 20-plus certification programs to standards-based curriculum and assessments—including electronic portfolios—and the simultaneous running and funding of both old and new programs over three years. Last, we made it through an Illinois Board of Higher Education degree programs review; and will experience, also this year, a North Central Accreditation review.

Lesson learned? Thinking systemically and keeping our mind on the end goal sustained and continues to sustain us through these years of labor- and time-intensive program and accreditation reviews.

**Addressing the leaky pipeline: A math story.**

Initially we looked at who was in our teacher candidate pipeline at UIC. We were unpleasantly surprised to discover that we lost fully half our pipeline before they ever reached our college. More specifically, we were losing many prospective candidates who had graduated from Chicago Public Schools, including nearly
everysingle African American student and most Latino students, not to mention a class salutatory or two. When we learned how to access students’ grades, we saw immediately that we were losing our students to remedial mathematics courses. We lost them from both the lower and the more advanced remedial course. We lost students who had placed out of the remedial course but who accidentally made their way back into the course. We found students who had received multiple failing grades for taking the same class over and over, before dropping out. We found students who had accumulated nearly enough credit hours to graduate, but who had not yet applied to our college by their junior year.

We retrieved data on current and graduated students and carefully analyzed entering data like ACT scores and placement test scores against how well they were faring in the program. We learned that students who entered UIC needing to take the less advanced math course rarely made it through to graduation and subsequent certification. We negotiated with the Mathematics Department to pilot a summer course taken in lieu of the more advanced remedial course and discovered that most students came in with what a faculty member called “punctuated knowledge,” i.e., gaps in a mathematics foundation. Eventually, Mathematics allowed us to substitute this course for its remedial mathematics course; over the semesters, we learned that 79 percent of students taking our course passed the first required mathematics course for elementary teachers, while only 52 percent of those taking the traditional remedial course remained standing. We like to believe that our work influenced the current campus post-secondary collaboration with Chicago Public Schools.

Lesson learned? In the end, this meticulous course of actions enabled us to see through to teacher certification a significant number of CPS graduates, including Black and Latino students, who would otherwise not have become teachers.

Diversifying the portfolio to address teacher shortages: Alternative Certification 101.

Like most schools of education, we came late to the realization that we were contributing to the oversupply of elementary and English and history secondary teachers, while not preparing an adequate supply of mathematics and science teachers. Indeed, UIC had years earlier called a moratorium on the preparation of biology teachers, because the department head had no resources to spare once he took care of the hundreds of biology majors who aspired to medical school. When the then-Provost offered me an opportunity and seed funding to develop an alternative route to certification, we seized the moment to draft and secure ISBE approval for a middle grades science and a middle grades mathematics program. Thus began four years of math and science teacher cohorts whose members comprised what one CPS administrator referred to as “shining stars” in the district. While we ultimately shut down the programs because the district cut its contribution to $2,000 a teacher for a program which cost $18,000 a teacher (we subsidized $6,000 a teacher through grants), and because No Child Left Behind (NCLB) adversely affected the hiring of our teachers, we learned enormous lessons from the programs. Among the lessons, we realized the extraordinary faculty dedication and resources it took to prepare quality teachers to dive into the everyday rigors of teaching preadolescents mathematics and science in under-served urban schools, including how our standard curriculum needed to be modified and rearranged to accommodate teachers’ day-to-day needs. We learned that, whereas 4 to 8 observations in 15 weeks sufficed for student teachers in traditional teacher preparation programs, once weekly was often insufficient to support alternative certification candidates on the job. We learned that teachers in shortage areas rapidly rise to department chair and other leadership positions. And, now that the programs are no more, faculty have incorporated lessons learned from the experience into our regular teacher preparation programs.
Lesson learned? From our experience with alternative certification, we learned that we can sometimes achieve our objectives by taking a course of lesser resistance (and, in the process, learn how to do what we were doing even better).

The teacher preparation curriculum and the issue of faculty governance.

The faculty are quite clear on one thing: Curriculum is the exclusive province of faculty, and no dean is supposed to be anywhere near it. Despite acknowledging this (grudgingly), I have been able to influence the curriculum in at least three ways. First is by “going with the goers,” via moral and on occasion financial support for initiatives. I gladly contributed supplementary funding to a joint Arts and Sciences/Education community college collaborative which developed, implemented, and evaluated a superb sequence of natural science courses for non-science and elementary majors. The relationship of many years has spun off many related curricula, supported a joint hire with Chemistry, and brought full-time predoctoral fellows to our professional development school. I nursed along small curricular efforts in mathematics, until a critical mass of mathematics educators emerged, and then I lent wholehearted support.

Second is by stimulating faculty dialogue via provocative seminars and gauntlet-tossing challenges. We have gathered together all of the alternative certification providers in Chicago. We have brought together all of the UIC faculty associated with any CPS K-12 mathematics program or initiative. We have invited in national experts to tell us what they think we should be thinking about. I have afforded all of our educational psychology faculty opportunities to think together about what makes their development and learning curriculum useful to prospective urban teachers. Similarly, our social foundations faculty have thought about what their curriculum offers prospective urban teachers beyond critique.

Third, most importantly, I have been able to influence teacher preparation curricula by the faculty hires I have made, although it took me several years to fully appreciate this power. I used precious grant resources to leverage yet more precious state resources to hire a critical mass of urban teacher educators, including Black and Latino faculty—not as many as originally promised, but important nonetheless. It has been a huge challenge retaining faculty of color, incidentally, because they are constantly being recruited by other institutions; yet they are a critical voice in our evolving teacher preparation curricula. We have also invested in hiring and developing a promotion ladder for full-time clinical faculty members whom we see as essential players in developing and sustaining relationships with cooperating teachers and schools where turnover is high. Despite their value, we have been unfortunately unable to pay clinical faculty at the rates they would command were they to return to teaching in K-12 school systems.

We have been able to parlay faculty hiring opportunities into joint Arts and Sciences/Education positions. We now have joint faculty in Mathematics (2), Chemistry (1), English (1), and Psychology/Learning Sciences (2). We shared a faculty member with Spanish until she decided to forego her program coordinator position, and we participated in the hire of a history teacher educator.

Lesson learned? While curriculum is the province of faculty, the dean hires the faculty.

Teacher professional development as a means to develop student teaching sites.

Early in the 10 years, we created an in-house database to track all of the field placements we had made for our teacher candidates. Soon afterwards, the Illinois Teacher Data Warehouse was developed in collaboration with ISBE, and schools of education were shortly able to access their own data. The ITDW is able to tag the school locations of every job our teacher candidates assume in Illinois public schools. We eagerly accessed our data, mapped
the data onto a map of Chicago schools, and immediately realized to our great chagrin that almost every placement we made for our teacher candidates was either in a magnet school or a school with a selective enrollment policy or in a predominantly Latino school. Very few of our placements were made in predominantly African American schools which comprise half of Chicago’s 600 public schools. Similarly, few candidates took their first teaching positions in CPS African American schools, although more taught in Black schools than were placed in Black schools; the vast majority of candidates taking CPS positions showed up in magnet or Latino schools.

I realized we needed to effect a redistribution of UIC teachers in CPS schools serving African American communities, if we were to be part of the solution of turning around low-performing schools. I deliberately targeted Chicago’s Near West Side, because another unit at UIC had a Neighborhoods Initiative associated with the community. We later included the North Lawndale neighborhood at a community partner’s request, thus achieving a partnership of sorts with schools with some of the highest teacher turnover rates in the district.

Getting faculty to place their students in our new partner schools was no easy matter. Most faculty had long ago dispensed with worries about school violence and neighborhood safety concerns, but here I was asking faculty to forego their long-established relationships with cooperating teachers and principals in familiar clinical sites, and to transplant their student teachers in sites some viewed outright as inferior placements—places where faculty perceived our students to be “hijacked” by school cultures which embraced different philosophies and pedagogies than ones they embraced at UIC.

Lessons abound. We learned that offering significant professional development workshops and opportunities for teachers in partner schools over several years oftentimes built trust and sometimes skill to the point where teachers gladly served as cooperating teachers for student teachers. When things go well, student teachers receive job offers in the schools where they student taught. On other occasions, we unanimously agreed that a particular school was not a good learning environment for student teachers. On still other occasions, we disagree with one another about whether the teaching we’re observing is effective teaching or not, and we can see that we are making our evaluations through our respective cultural lenses.

Overall lesson learned? Clearly, the school site is itself a launch pad not only for teacher professional development, but also a site for nurturing student teachers and fostering professional conversations among teacher educators.

**Doing the best we can by our new teachers.**

It has been challenging to determine the best way to support our teachers post-graduation. Initially, several faculty members devoted considerable energy and resources to collaborating with the district on a system-wide new teacher induction and mentoring program. As soon as UIC left the partnership, the program faltered; the Department of Professional Development was moved under Human Resources, staff numbers were decimated by budget cuts, and only the shell of a once-robust program remains. While we continue to partner with the district on new teacher support via a Teacher Quality Enhancement grant, the primary relationships and communication networks with new teachers are sustained with UIC program faculty.

Three years ago, we were the beneficiaries of a hefty grant to support new teachers in West Side schools with in-classroom support. Not only did we realize that this was a non-replicable (but extremely welcomed) program but, much to our dismay, new UIC teachers last hired by West Side partnership principals were the first to be let go in last year’s round of budget cuts. In the interest of developing better job protections for our graduates in CPS schools, we are talking with our school leader program coordinator about the possibility of placing teachers with our own principals.

Lesson learned? This one is a work in progress.
Strategies for changing business as usual.

Unquestionably, investing in data—including database development and a person to track data—has been essential to making the case for change in business as usual and in soliciting buy-in. Without looking at data which show our lack of presence in African American schools, I could not have made it our moral obligation to change our practices. Without seeing where our students sat in the preparation pipeline, we would not have thought to take steps to address mathematics gaps. Without annually capturing all of our teacher professional development activities, we might not have noticed that these relationships were yielding student teaching opportunities and thus begun to deliberately cultivate these opportunities. Without looking at data, we would not have learned that a disproportionate number of our math concentrators were electing to student teach in primary grades, rather than the middle grades for which they were so supremely prepared. And so on. Data enable us to focus on what matters most in pipelining our teacher candidates to Chicago schools where they are most needed.

Securing external funding, too, was absolutely critical in getting started in changing business as usual, especially as public universities fell victim to budget droughts. The original large, multi-year grant from the MacArthur Foundation was pivotal in leveraging state dollars for faculty hires, as well as for leveraging smaller foundation grants to support associated components of the program. In addition, the 18-month lead-up to actually securing the grant forced collaborative planning which might otherwise not have happened.

It is necessary to comment on the downside of a dean securing grants when trying to change the culture of a college. One has to walk a fine line between supporting faculty efforts, on the one hand, and pressuring faculty to realize the dean’s vision, on the other. In the end, most would agree that having resources is much better than having no resources.

A third “strategy” worth mentioning is the habit of capitalizing upon opportunities, if they appear to support the larger agenda of supplying teachers to low-performing schools. Two and a half years ago, we were invited to serve as contractor for a professional development school serving the children residing in two Chicago public housing communities. Three principals later, it has been and continues to be the wildest and most enlightening of immersion experiences. No experience before or since this one has taught me and my colleagues so much about how a large urban school bureaucracy influences the life of its constituencies, including its university partner.

Lesson learned? You can never predict when you’re going to learn something, but data and money assist immeasurably.

Concluding remarks.

The above comments provide a glimpse into some of the details of our decade-long culture change to relieve some of the teacher shortages of Chicago Public Schools. Have we been successful? By many measures, I believe so. For most of the 10 years, despite our relatively small size, we have been the top supplier of new teachers to CPS, and the top supplier of teachers overall. A neighborhood association’s report of recent hires in North Lawndale schools showed that the number of new hires who graduated from UIC was double the total number of new hires from all other colleges and universities combined. Principals regularly ask us for our graduates, and tell us how much they value UIC teachers in their schools. Many more secondary teachers are taking positions in CPS high schools, including mathematics teachers. Our ITDW tracking shows us that our teachers tend to stay in Chicago public schools, once hired, and tend to stay either in their original schools or in schools with similar racial-ethnic characteristics. Our four-year old Ed.D. school leadership program already boasts 17 principals, 13 assistant principals, 2 system-level administrators, and several principal interns.
During this period, UIC faculty have served as directors of the district’s reading initiative (2) and the district’s math and science initiative, as well as one of two lead partners on the district’s mid-tier reading initiative. Although we have bumps and bruises from the experiences, sometimes from our own internal struggles about courses of action, we have more staying power with the district than most higher education institutions and the journey has been well worth the while.
Can the No Child Left Behind Act Solve the Problem of Underqualified Teachers?

Richard M. Ingersoll, Ph.D.
Professor of Education and Sociology
University of Pennsylvania

Few education problems have received more attention in recent years than the failure to ensure the nation’s classrooms are staffed by qualified teachers. Over the past decade, dozens of studies, commissions and national reports have bemoaned the qualifications and quality of our teachers. As a result, reformers in many states have pushed for more rigorous teacher education and licensing standards. Moreover, a whole host of initiatives and programs have sprung up which are designed to recruit new and talented candidates into teaching. Concern with the quality and qualifications of teachers is neither unique nor surprising. Elementary and secondary schooling are mandated in the U.S. and it is into the care of teachers that children are legally placed for a significant portion of their lives. The quality of teachers and teaching are undoubtedly among the most important factors shaping the learning and growth of students. Moreover, the largest single component of the cost of education in any country is teacher compensation.

Over the past decade and a half, I have been doing research on the problem of under-qualified teachers. I have come to the conclusion that, while staffing the nation’s classrooms with qualified teachers is one of the most important problems facing schools, it is also one of the most misunderstood. This misunderstanding centers on the supposed sources of the problem—the reasons behind the problem of underqualified teachers in American schools. I have come to the conclusion that, unfortunately, most recent teacher-reform efforts will not solve the problem because they do not address some of the key causes.

The No Child Left Behind Act

The most significant contemporary education reform effort is the federal No Child Left Behind Act (NCLB)—the Elementary and Secondary Education Act revised and re-authorized in January 2002. This legislation set a new, unprecedented and laudable goal—to ensure that the nation’s public elementary and secondary students are all taught their core academic subjects by highly qualified teachers. According to NCLB, by the end of the 2006-2007 school year, all schools will have to meet this goal.

One of the difficulties encountered in such legislation is a lack of consensus about standards for defining a qualified teacher. Few would argue that teachers ought not to be qualified, and there is substantial research support that teacher qualifications are tied to student achievement. But, there has been a great deal of controversy concerning what kinds of preparation and requirements are necessary to be considered qualified. After much debate, NCLB defined a “highly qualified” teacher as someone who has a college bachelor’s degree, who holds a full state-approved teaching certificate (or license), and who is competent in each of the core academic fields they teach. There
are several means by which teachers can establish “competency” in a field. They can hold an undergraduate or graduate major or its equivalent in the field, pass a test in the field, hold an advanced teaching certificate in the field, or meet some other state-approved method of evaluation for the field.

In order to assess how well schools are doing in regard to these new requirements, and to hold them accountable if they do not meet them, NCLB requires four things of states and school districts: annual report cards on teacher qualifications for every school; plans for improvement; disclosure of teachers’ qualifications to the public upon request; and a letter home to parents of students taught by underqualified teachers.

**What the Data Tell Us About the Problem**

The most comprehensive and accurate data on teachers’ qualifications come from the U.S. Department of Education’s *Schools and Staffing Survey*. This is a very large survey administered nationwide to teachers every few years since the late 1980s. What do these data tell us?

The data indicate that most teachers have a basic college education and full teacher certification. In the 1999-2000 school year, just prior to the enactment of NCLB, 99 percent of public school teachers in the United States held a bachelor’s degree, and almost half held a master’s degree or higher. Nearly 92 percent of teachers in public schools held full, regular teaching certificates. This most recent cycle of the survey was conducted in 2003-2004, a year after the enactment of NCLB. The results changed very little. Again, 99 percent of public school teachers in the United States held a bachelor’s degree, and almost half held a graduate degree. Again, nearly 92 percent in public schools held full, regular teaching certificates.

These data appear to conflict with conventional wisdom. In recent years, there have been numerous reports that many school districts—especially those serving low-income, urban communities—have been forced to hire significant numbers of uncertified teachers to fill their teaching vacancies. This is no doubt true in some districts. The national data suggest, however, that overall the number of teachers without a full certificate comprise only a small proportion of the K-12 public teaching force. In a workforce of almost 3.25 million teachers, about 235,000 teachers have a less-than-full certificate and about 48,000 teachers have no certificate at all. (Less-than-full refers to certificates issued on a temporary basis to allow those who have not yet completed all of the requirements to begin teaching immediately).

Of course, having these credentials is no guarantee of quality, or of competency in a given field. But, contrary to conventional wisdom, at first glance it does appear that the vast majority of teachers meet the highly qualified teacher requirements of NCLB and, indeed, that is exactly what many states have been declaring in their annual reports.

However, this is only part of the story the data tell. The data also show large numbers of classrooms are staffed by teachers who are not qualified in the subject taught, and hence, do not meet the NCLB standard. This is the result of a little-known, but widespread, practice called out-of-field teaching—teachers assigned subjects that do not match their training or education. Highly qualified teachers may become highly unqualified if assigned to teach subjects for which they have little preparation. Those trained, for instance, in how to teach social studies are unlikely to have a solid understanding of how to teach biology or reading; a math teacher may not have sufficient background to teach biology. This problem has also been recognized by NCLB, which explicitly defines highly qualified teachers as those who are competent in each of the academic subjects they are assigned to teach.

The data from the 1999-2000 school year indicate that, at the secondary school level, about one-fifth of classes in each of the core academic subjects (math, science, English, social studies) have teachers lacking full teaching certificates in the subject taught. Some
schools and classes have more out-of-field teaching than others. Within schools, lower-ability level classes more often have out-of-field teachers than do gifted, advanced and college preparatory classes. High-poverty schools have more out-of-field teaching than more affluent schools. Indeed, while teachers in disadvantaged schools lag only slightly behind teachers in other schools in their formal qualifications, they are far more likely to be assigned to teach out of field.

The most recent national data—from the 2003-2004 school year—have been recently released and can be used to show if there has been any change in levels of out-of-field teaching since the enactment of NCLB. We are currently in the process of analyzing the newer data toward that objective, but have not yet completed our calculations. However, in the preliminary analyses thus far it appears that overall there was little change in out-of-field teaching in the year after the advent of NCLB.

The Sources of the Problem

The crucial question, and the source of great misunderstanding, is why so many classes are taught by teachers who are not qualified in the subject taught. Two interrelated explanations typically are advanced. The first holds that out-of-field teaching is a problem of poorly prepared teachers. Proponents typically argue that teacher preparation programs are overloaded with required courses in pedagogy and teaching methods; hence, the remedy is requiring prospective teachers to complete “real” undergraduate majors in academic disciplines.

Certainly there are some problems with the rigor and quality of teacher preparation standards and programs. But this explanation confounds two different sources of the problem of underqualified teachers by confusing teachers’ education prior to employment with teachers’ assignments after employed. The primary source of out-of-field teaching lies not in the education or training teachers have, but in the lack of fit between their preparation and their teaching assignments. Principals assign many teachers to instruct classes that do not match the field of their degree or certification, or both.

The key question thus becomes: why are so many teachers mis-assigned? In answer, many commentators turn to the explanation of teacher shortages. Shortfalls in numbers of available teachers, this argument holds, have forced many schools to assign teachers out of their fields.

School staffing and hiring difficulties are a factor in the degree of teacher mis-assignment, but the data reveal two important shortcomings with this explanation. First, it cannot explain high levels of out-of-field teaching in fields such as English and social studies, where there are surpluses of qualified teachers. Second, in any given field only a minority of schools report difficulty finding qualified candidates to fill their teaching job openings, and the data show that about half of all mis-assigned teachers work in schools that reported no such difficulties.

The data point toward another explanation: the manner in which schools are organized and teachers are managed. School staffing decisions usually are the prerogative of school principals, and teachers typically have little say over their assignments. School principals face the difficult task of providing an increasingly broad array of programs with limited resources, time, budgets, and teaching staff. But, within those constraints, principals have an unusual degree of discretion, and there is little centralized regulation of how teachers are utilized once they are hired. In this context, principals report that, from a managerial perspective, they find that assigning teachers to teach out of their fields is often more convenient, less expensive, and less time consuming than the alternatives.

For example, rather than hire a new part-time science teacher for two classes of a newly state-mandated science curriculum, a principal may find it simpler to assign two English or social studies teachers to cover those classes. If a teacher unexpectedly leaves mid-semester, a principal may find it faster and cheaper to hire
a readily-available, but not fully qualified, substitute teacher, rather than conduct a formal search for a new teacher. When faced with a tough choice between hiring an unqualified candidate for a math teacher position or doubling the class size of one of the fully qualified math teachers, a principal might opt for the former. From a managerial perspective, these choices may save time and money for the school, and ultimately for the taxpayer. From an educational perspective they are not cost free, as they are among the largest sources of under-qualified teachers in schools, now deemed to be illegal under NCLB.

The comparison with traditional professions is stark. Few would ask cardiologists to deliver babies, real estate lawyers to defend criminal cases, chemical engineers to design bridges, or sociology professors to teach English. Likewise for the high-skill blue-collar occupations; few would hire an electrician to solve a plumbing problem. The commonly-held assumption is that these lines of work require substantial expertise and, therefore, specialization is necessary. Despite research establishing that quality elementary and secondary teaching is highly complex work, K-12 teaching lacks commensurate stature and standing. Underlying out-of-field teaching, it appears, is the assumption that teaching requires far less skill, training, and expertise than many other kinds of work, and it is therefore acceptable to treat teachers like interchangeable blocks.

What is to be done?

Understanding the reasons behind the problem of underqualified teaching is important because of their implications for solving the problem. Most recent federal, state and local teacher policies and initiatives have focused on two general approaches to ensure that all classrooms are staffed with qualified teachers: upgrading the qualifications of teachers, and increasing the quantity of teachers supplied. These kinds of initiatives are also emphasized in NCLB. The Title II portion of the Act, for example, specifically focuses on enhancing teacher training and teacher recruitment in its list of methods approved for funding.

Underlying these kinds of methods is what might be called a teacher deficit perspective—the assumption supporting this perspective is that the primary source of underqualified teachers in schools lies in deficits in teachers themselves—their numbers, preparation, knowledge, ability, and licensing, etc. Of course, increasing teacher recruitment, and upgrading teacher preparation standards can be useful first steps. But, the above methods do not address the ways schools themselves contribute to the problem of underqualified teachers. The data tell us that solutions to the problem of out-of-field teaching must also look to how schools are managed and how teachers are utilized once on the job. In short, recruiting thousands of new candidates and providing them with rigorous preparation will not solve the problem if large numbers of such teachers continue to be assigned to teach subjects other than those for which they were prepared.

This problem presents a thorny policy dilemma at the federal level. Solving the problem of underqualified teachers requires changes to the internal management of schools. But the internal operation of schools has traditionally been a local affair and off limits to federal intervention. Moreover, if legislation, such as NCLB, results in increased accountability for teachers without commensurate changes in the way schools are managed, it could lead to a classic organizational problem—employees blamed for things over which they have no control—and end up exacerbating the very teacher quality problems the legislation seeks to solve.

If assigning teachers to teach out of their fields has been a prevalent administrative practice for decades because it is more efficient and less expensive than the alternatives, then its elimination will not be easily accomplished simply by legislative fiat. In order to meet the goal of ensuring all students are provided with qualified teachers, it will be necessary to rethink how districts and schools go about managing their human resources.
One area that will need rethinking is how school staffing decisions are made and who makes them. As mentioned, the data tell us that, unlike in higher education, elementary and secondary teacher staffing decisions have traditionally followed a top-down model: school principals make such decisions and teachers typically have little say over which courses they are assigned to teach. As the earlier examples illustrate, these staffing decisions often involve difficult tradeoffs and sometimes lead to out-of-field teaching. In contrast, districts and schools could implement mechanisms of school-based management where such decisions are shared with those who must live with, and may be held accountable for, the consequences—the faculty. Similarly, states could provide training and assistance to district and school administrators in how to better balance tradeoffs between organizational, budgetary and educational needs—something that has been largely neglected in the past.

Another area that will need rethinking concerns teacher employment practices. Meeting standards for qualified teachers will be more difficult in some settings than others. Rural school districts, for example, tend to have smaller secondary schools with smaller faculties. As a result, the data suggest, teachers in these schools are more often required to be generalists—teaching a variety of subjects regardless of their background. In these kinds of settings, states might consider the use of itinerant teachers, where schools could share the use of teachers with preparation in a specialty. This could include the employment of retired teachers. Similarly, states could fund technology in order to provide rural and hard-to-staff schools with access to teachers with preparation in a specialty.

A third area that will need rethinking concerns the provision of administrative support for teachers. The data indicate that beginning teachers are more likely than veteran teachers to be given out-of-field assignments. Disproportionately burdening newcomers probably contributes to the problem of high levels of beginning teacher attrition. Moreover, the data indicate that when teachers are mis-assigned they are largely left to their own devices. Where it is difficult to entirely eliminate out-of-field teaching, districts could prohibit out-of-field assignments for new teachers, provide funding for additional course work for mis-assigned teachers, or provide funding for veteran teachers to mentor, assist or team teach with mis-assigned teachers.

The lesson is clear: if we want to ensure that all classrooms are staffed by qualified teachers, we will need to change the way that schools have been operated and that teachers have been managed. In short, upgrading the quality of teaching will require upgrading the quality of the teaching job.
The Challenge of Special Populations to Accountability for All

Martha L. Thurlow, Ph.D.
Director, National Center on Educational Outcomes
University of Minnesota

Special populations are part of the U.S. educational system and a recognized part of No Child Left Behind (NCLB). Dramatic benefits to students with disabilities and English language learners have resulted from the implementation of NCLB. The challenge of special populations is recognizing that they can learn and that by giving them access to the general curriculum, along with the accommodations, assistive technology, supports and services that they need, they too can fulfill the vision of “no child left behind.” Those who identify other challenges often are starting from misperceptions about who the students are, their capabilities for learning, or whether educators should be held responsible for ensuring that they learn.

NCLB has done more to bring attention to students with disabilities, their need for access to the curriculum, for assistive technology, for individualized services, supports, and specialized instruction, and for better assessments, than previous laws and requirements. Those who work with English language learners will echo this statement. Despite the fact that it may be challenging for schools to rethink traditional but ineffective practices and separate curricula for students with disabilities or English learners, NCLB has been good for students with disabilities and for students who are learning English. For students with disabilities in particular, NCLB has ensured that they are finally being given access to the curriculum that the Individuals With Disabilities Education Act required that they be given access to since 1997.

Stories in the media do not necessarily paint the same picture that I will portray. That is because there remains a tension between the views of those being held responsible for implementation—administrators and educators on the one hand, and students and their parents on the other hand. Frequently the stories in the news reflect the perspectives of those who do not want to be held responsible for students with disabilities, whom they still view as the most difficult to teach of their students, and in many cases, whom they view as those students who cannot learn. This is not the case everywhere, and I will show that where educators believe that students with disabilities can learn, the mandate of NCLB has been taken seriously and students with disabilities have made tremendous improvements in performance, with the achievement gap between students with and without disabilities narrowing to an impressive extent.

NCLB’s vision of accountability for all, and its goal to leave no child behind, has been especially challenging for the special education subgroup—to the point that districts and states have devised ways not to be accountable for some of these children, and to the point that the U.S. Department of Education has felt the need to devise “flexibilities” for these students beyond those for the students with the most significant cognitive disabilities and beyond those
for all other students. Because of the pervasive difficulties facing the inclusion of students with disabilities in NCLB, I will focus my remarks on this subgroup.

Before describing the benefits and challenges of NCLB, there are two important basic pieces of information that should be revisited: first, a quick reminder of why state NCLB assessments are appropriate for students with disabilities; and, second, a refresher on who students with disabilities are so that we all are thinking about the same students.

Some Basics

Why NCLB Assessments are Appropriate. A primary function of NCLB assessments is to give a picture of achievement, and to serve as a warning flag—something that alerts us to review low achievement results, so that something can be done about it. We need to know when students are not performing well. This is particularly important for students with disabilities, who in many schools have not had meaningful access to the general education curriculum. In other words, many students with disabilities have not received the accommodations, services, supports, and specialized instruction they need to succeed in the same challenging curriculum as their peers, as required by the IDEA. Assessments required by NCLB and IDEA can provide a window to shed light on whether meaningful access and high quality instruction has been provided, but assessments can improve outcomes only if limited access or poor instruction change as a result of that view.

Who Are Students with Disabilities? It is also important to remember who students with disabilities are. There are misperceptions among the public, policymakers, and even educators about the makeup of this subgroup. Seventy-five percent of students in special education have learning disabilities, speech/language impairments, and emotional/behavioral disabilities. These students, along with those who have physical, visual, hearing, and other health impairments (another 4-5%), are all students without intellectual impairments. When given appropriate accommodations, services, supports, and specialized instruction, these students (totaling about 80% of students with disabilities) can learn the grade-level content in the general education curriculum by going around the effects of their disabilities, and thus achieve proficiency on the grade-level content standards. In addition, research suggests that many of the small percentage of students with disabilities who have intellectual impairments (less than 2% of the total population of all students, which is less than 20% of all students with disabilities), can also achieve proficiency when they receive high quality instruction in the grade-level content, appropriate services and supports, and appropriate accommodations. An example of such a student—one who does have an intellectual disability but who with adequate school supports and services, and with accommodations during instruction and during assessment, was able to obtain access to the content—is Katie Bartlett, a young woman with Down syndrome, who passed the high school competency exam in Massachusetts. Katie was held to the same standards as other students and met them.

Expecting less than proficiency on grade level content from students like Katie Bartlett, and indeed expecting less from any student with a disability (or any English learner), reflects little understanding of who these children are. Students with disabilities can meet standards. Special education eligibility should not be an excuse to expect less from a child, nor to provide little to a child. The assessment and accountability provisions of NCLB have helped us recognize and shed light on the outrageous problem of some educators’ low expectations for these children, reflected in statements like this: “If special education students could perform well on these tests, they wouldn’t be in special education.”

What Have Been the Benefits of No Child Left Behind for Students with Disabilities?

Data are emerging that show increases in actual achievement and other effects as well. For schools to produce the increases in achieve-
ment, there has to first be the belief system that students with disabilities can learn and, second, a commitment to their inclusion in the accountability system. These two conditions, of course, relate to the challenges that continue to surround students with disabilities during the fifth year of the implementation of NCLB. Statements made in the news media by educators and administrators reveal that the belief system that students with disabilities can learn is not yet pervasive. Other indicators of questionable beliefs about the special education subgroup are the decisions of state policymakers to require a larger number of students in the disability subgroup than in other subgroups for results to be reported or included in accountability calculations; ongoing demands for flexibility for additional students; and a pattern of omission from key policy decisions, such as recently approved growth models that overlooked students with the most significant cognitive disabilities. Nevertheless, pockets of believers who have resisted loopholes in accountability for the disability subgroup are showing that NCLB data are important for students with disabilities and can serve exactly the purpose for which it was intended for this subgroup.

State directors of special education have responded to National Center on Educational Outcomes surveys about the improvements that they are seeing in the performance of their students, attributing the improvements to clear participation policies, alignment of Individual Education Programs with standards, improved professional development, development and provision of accommodation guidelines and training, increased access to standards-based instruction, and improved data collection (Thompson, Johnstone, Thurlow, & Altman, 2005). Analyses of publicly reported assessment data since 2000-2001 show improvements in the transparency of data for students with disabilities, both for participation and for performance (work in progress, 2006). For example, NCEO’s identification of states with clear participation reporting to the public for students with disabilities showed only 5 states in 2000-2001, but 20 states in 2004-2005. These data also showed large increases in participation percentages across time for most states. Data on performance showed similar changes—more states with clear transparent reporting, and increases in performance across years). Students in one state have moved from 30% of 10th graders attaining the proficiency designation in 2003 to 32% in 2004, to 46% in 2005, to 50% in 2006. This improvement is under the status model of improvement—different students each year. Data from the National Assessment of Educational Progress (NAEP) support these data—overall there has been an increase in the performance of students with disabilities in both reading and math in grade 4 each year since 1996—greater increases than for students without disabilities.

Students with disabilities are receiving more appropriate accommodations (such as large print editions, extra breaks, sign language interpreters, calculators, individual administrations) during testing. While we are still improving our knowledge base on the kinds of changes to the practices and procedures during testing that are appropriate and do not invalidate scores, and states are still determining their policies in line with what they are testing, all the attention on accommodations has lifted awareness in the field so that students who need them are more likely today than before NCLB to be receiving needed accommodations—during testing and during instruction. Prior to NCLB, high estimates indicated that the use of testing accommodations was about 53% for elementary schools and 44% for middle and high schools (Thompson & Thurlow, 1999). Today those percentages are 65% for elementary schools, 64% for middle schools, and 62% for high schools (Thurlow, Moen, & Altman, 2006). We do not know, as yet, what the percentages should be, but having them closer to the same across school levels is a good sign.

Studies are now being conducted to determine what produces good performance among students with disabilities. One of these studies, which was conducted in Massachusetts, identi-
fied low income urban schools where special education students were performing better than expected. The researchers then investigated what was happening in those schools for the special education students. They found 11 common factors that characterized those schools where students with disabilities were performing well: (1) a pervasive emphasis on the curriculum and alignment with the standards, (2) effective systems to support curriculum alignment, (3) emphasis on inclusion and access to the curriculum, (4) culture and practices that support high standards and student achievement, (5) well-disciplined academic and social environment, (6) use of student data to inform decision making, (7) unified practice supported by targeted professional development, (8) access to resources to support key initiatives, (9) effective staff recruitment, retention, and deployment, (10) flexible leaders and staff that work effectively in a dynamic environment, and (11) effective leadership (Donahue Institute, 2004).

One group of students for whom the NCLB requirements have had tremendous impact is students with significant cognitive disabilities, individuals who for the most part have been in the educational system only since the passage of Public Law 94-142, the Education of All Handicapped Children Act. Prior to this, some of these children would have been in institutions, or in care-giving facilities. In December 2003 the 1% rule permitted states to develop alternate assessments that allow proficiency to be defined differently from the regular assessment. Alternate assessments were first required in IDEA, but NCLB added accountability requirements. States could set alternate achievement standards on the results of these alternate assessments in order to define proficiency for students with the most significant cognitive disabilities. These assessments are identified as assessments for students who are unable to take regular assessments. In those states where the right students were targeted, these alternate assessments have resulted in a sea change to the instruction that students with the most significant cognitive disabilities are receiving. Assessment has changed their school experiences. They are now receiving an education—access to the general curriculum—rather than a separate curriculum, or even simply receiving care in the schools. NCEO has data from state assessments documenting the shift to the general curriculum, and qualitative data on the perceptions of educators on the benefit of this shift (Moore-Lamminen & Olsen, 2005). Impact data are being collected nationally by the National Alternate Assessment Center.

Attempts are being made to collect information on overall impacts of the accountability requirements. One of these is time spent in general education environments, which is a proxy for access to the general curriculum. The just released 26th Annual Report to Congress reported that the percentage of time spent in the general education classroom had increased for students with disabilities through the 2002 data available in the report. The 2005 NCEO survey of state directors of special education suggested that NCLB continues to expand this access to the general curriculum for students with disabilities.

What Have Been the Benefits of No Child Left Behind for Assessments?

There have been several benefits to assessment systems that have resulted from requirements to include all students, particularly those with disabilities and English learners. Assessment developers who have typically been able to exclude whoever did not fit the norm have had to revisit assumptions and revise assessment models. This has been good for the assessment world. States, in particular, have focused on how to make their assessments more accessible—through better accommodation policies and by applying universal design principles to assessments—not by changing the construct that is to be measured or reducing the grade-level content that is being assessed. Rather, developers have focused on ensuring that tests really measure what they are intended to measure—not extraneous factors, such as whether the student can figure out what the test devel-
oper means by a question or whether a picture has important clues about the answer to a question. Identifying ways to improve assessments for students with disabilities has, in fact, resulted in improving assessments for all students.

What Have Been the Challenges of No Child Left Behind for Accountability?

Despite the many benefits of No Child Left Behind, there has been an undercurrent of challenges for the disability subgroup. These have manifested themselves in a number of ways. The most obvious was the decision made early on in many states that assessment results would be publicly reported and included in accountability for students in the disability subgroup only when a larger number was included relative to those in other subgroups. It is not uncommon to find that reporting and accountability occurred in a school, for example, for all subgroups except the disability subgroup when there were 30 students in a group; for students with disabilities, the number had be 45 students. Despite detailed arguments about error rates and confidentiality, it seemed rather to be a policy decision—a way to avoid responsibility for a group perceived to be the lowest performing subgroup. Stories also soon spread about districts having students disappear from the denominator by moving them around, either from school to school, or from one grade to another, or by having their scores count for participation but not for performance. We heard about the “EGG” Game being played by principals—the “enrolled grade game”—holding 9th grade students back one year, then promoting them to 11th grade to ensure they did not take the 10th grade test, which was the accountability test. Not all states and districts are doing these things, and some states have policies and monitoring to prevent it. When they do these “games,” it hides the very thing that we need to know about—the performance of low performing students—very often, students with disabilities.

The overall initial poor performance of students in the disability subgroup may be the reason that states, districts, and schools have reacted out of fear and have tried to identify loopholes for their accountability systems. It also may be the reason that they and others are actively seeking other loopholes and additional flexibility in a system that is designed to do exactly what it should do—hold the system accountable for students who are not reaching grade-level achievement standards.

Where Should No Child Left Behind Go From Here?

I can remember the time before NCLB and before IDEA in 1997, when there were no requirements for students with disabilities (nor for students with limited English proficiency for that matter) to be included in the assessment systems along with other students. I recall stories of these students being sent on field trips to the zoo, or the parents of these students being told kindly by the teacher or the school principal that their students could stay home from school the day that the state test was given. I remember watching the referral rates of students to special education creep up not so slowly as more and more students were exempted from school testing in those states that had accountability systems in place and did not want to be accountable for all of their students. Their solution was to get rid of the students for whom they did not want to be accountable.

The solution is NOT to get rid of students who are tough to teach. It was not the correct solution then, and it is not the correct solution now. We have enough emerging evidence that students with disabilities can learn when given access to the curriculum and rich content instruction by qualified teachers that we can no longer justify coming up with a different accountability system for them. There are some principles that need to guide our approach to accountability for all—to ensure that students with disabilities thrive along with other students.

1. We need to recognize that all students with disabilities are general education students first—they need access to the general curriculum, qualified teachers, and high
expectations for their learning. With accommodations, services, and supports, we should expect 85-90% of these students to achieve grade level mastery on achievement tests that are administered in today’s schools. The remaining students have significant cognitive disabilities, and they achieve proficiency through the appropriate option of alternate achievement standards as provided for by the 1% regulation. Thus, we need to treat students with disabilities the same as other students. Except for the 1%, and the recognition of the need for accommodations, services, and supports as provided for by the IEP, the same approaches should exist for students with disabilities as exist for general education students.

2. We need to focus attention on the lowest performing students, not try to hide their performance or get them out of the system. Experiences with the 1% students have shown that there have been dramatic benefits to this population through the requirement that they be assessed, and for their performance to be reported separately. A strong research effort has developed around this type of assessment, and there are now systematic methods to ensure greater access to the general curriculum for these students. These are succeeding both in ensuring access to the curriculum and in increasing student performance to levels educators and parents thought impossible. There are many other students with disabilities who are performing at very low levels—students with and without intellectual disabilities. We need to focus on these students, the lowest performing students—to open up reporting and provide information on what is happening instructionally, and in the provision of accommodations and services. Based on previous research done in some states, I am confident that by looking at the lowest performing students on the regular assessment, for example the lowest 20%, a state will find not only students with disabilities, but low performing general education students as well. Then, the state can focus more appropriately on the issues that exist for all of those students—access to the curriculum, improved instruction, and yes, accommodations, and special education services as well if appropriate.

3. Adjustments to the accountability system itself should be made with care and should apply to all students, not to one subgroup. The only adjustments needed for the disability subgroup have been made—the 1% flexibility, the recognition that tests should be universally designed (which really benefits all students), the requirement that accommodations should be provided to those students needing them, and separate reporting by subgroup. Other adjustments to accountability systems should be made for all students, not just one subgroup, with consideration of intended and unintended consequences for students overall and for student subgroups. Moving to an accountability system based solely on growth unrelated to absolute standards, e.g., individualized growth set by an IEP, is an example of an accountability system that would have unintended negative consequences for the disability subgroup (and others as well). Students with disabilities for years have shown growth against IEP goals.

The critical difference that NCLB is making for these students is the link to grade-level content standards, a link that will bring them up to the level that they need to meet to be successful to move forward in their school work and to become contributing members of a competitive global society.
Introduction

The combination of the Adequate Yearly Progress requirements in No Child Left Behind and state results on the National Assessment of Education Progress have fueled renewed attention to the idea of national standards. NCLB has given added importance to how each state sets the bar for proficient on its own tests, while the persistent gaps between the percent of students scoring proficient on state tests compared to the percent scoring proficient on state NAEP highlights the enormous variability among states. Yet being proficient in reading by the 4th grade shouldn’t vary from state to state. We need consistent nationwide standards because the knowledge and skills young people need in core academic subjects such as reading and math is neither defined nor constrained by state boundaries.

But national standards alone won’t cure what ails 50 state standards. More important than the inconsistency among state standards is the fact that state standards themselves are not aligned with the knowledge and skills students must have in order to succeed after they leave high school, particularly in postsecondary education and the workplace. Simply put, today’s state standards—as well as the national standards developed in the late 1980s and early 1990s—reflect a consensus among subject matter experts about what would be desirable or even important for young people to learn. They are not the result of a careful analysis of the work young people will do when they complete K-12 education, and the knowledge and skills essential for postsecondary success. Consequently many students in every state meet state standards, pass state tests, and complete state-required courses, only to require remedial courses once they enroll in college. They may have been proficient, but were obviously not prepared.

Whether or not “national standards” is an appropriate response to these circumstances depends entirely on what is meant by national standards and how they are approached. Efforts to create national standards in the past—from President George H. W. Bush’s America 2000 program and American Achievement Exams to President Clinton’s Goals 2000 and Voluntary National Test initiatives—have produced considerably more political heat than education light;
these battles diverted attention away from important work that could have led to greater progress. In hindsight these initiatives also suffered from significant design flaws that would have limited their effectiveness absent the political obstacles they encountered. Any attempt to set national standards now should draw on the lessons from these previous efforts. We need to avoid rekindling the same political battles, and rethink the design and approach.

**Lessons From Federal Efforts**

Though they differed in important respects, these initiatives had several things in common. They were federally financed and directed. The standards and tests were for the voluntary use by states. The Bush Administration funded a set of organizations picked noncompetitively to develop national standards in the core academic subjects—including the highly controversial U.S. History Standards and the widely criticized English Language Arts Standards. The standards were either judged in the political arena or by Education Department staff monitoring the grant; there was no effective mechanism for independent quality review of the standards developed by the pre-selected organizations. The Clinton Administration’s Goals 2000 legislation authorized the creation of a National Education Standards and Improvement Council, intended to provide a quality control mechanism for the national standards project. However, fears of a “national school board” that would dictate local curriculum from Washington rendered the council so controversial that the provision was repealed by Congress before the members of the council could be appointed. Clinton’s second attempt at national standards and tests were the Voluntary National Tests in 4th grade reading and 8th grade math. Based on the highly regarded NAEP frameworks, these were to be overseen by the National Assessment Governing Board, whose members are appointed by the Secretary of Education. This proposal too proved to be highly controversial.

None of these initiatives accomplished what they set out to do. These efforts failed because collectively they suffered from three key weaknesses.

They provoked heated debates about who is in charge. In different ways each of these initiatives provoked debates over which level of government ought to be responsible for education standards and curriculum, and over the ideological persuasion of the individuals or organizations involved in the process. These fights prevented almost all of the federal initiatives from enactment and/or implementation. But they did more than that—the debates often eroded public support for standards-based reform led by any level of government, started ideological and often partisan battles that spilled over to the state and local levels, and distracted attention that could otherwise have helped standards based reform move forward. As these developments played out, state officials were particularly cautious about advancing standards-based reform, and were often forced to spend more time and attention protecting their efforts from political attack than ensuring their educational quality.

These fights also obscured a much larger point. In reality, neither the federal government nor state and local education agencies set standards. The real world that students enter after they complete high school, particularly colleges, employers and the military, are setting the standards students must be prepared to meet. Effective standard-setting efforts, at whatever level of government, will translate the signals these institutions send into the education language of standards and curriculum in order to inform what goes on inside schools.

They failed to anchor standards in the real-world demands students face—and should be prepared for—after completing high school. The impetus for the standards movement was in part to raise expectations so that young people could compete in a high-skill knowledge-based economy. But the voluntary national standards that were created in the early 1990s represented a consensus among experts in each discipline as to what is important for young people to learn by the time they complete high school, rather than emerging from a careful analysis of the
skills needed for success after high school. Building consensus, particularly among a diverse group of content experts, is important. But it fueled the political nature of the standards process, resulting in standards that were often a mushy compromise among competing views, or vulnerable to attack by one end of the political spectrum or the other. And by failing to anchor the standards in real world demands, the authors and promoters of standards denied themselves a powerful source of legitimacy.

They did not promote a coherent approach to standards based reform. Improving instruction and achievement requires a coherent approach to standards-based reform. This means that summative and formative assessments, curriculum, instructional strategies and materials all need to be highly aligned with academic standards. With the benefit of hindsight, it is clear that none of the previous federal efforts were adequately designed to promote needed coherence, and probably couldn’t be given the political obstacles mentioned above.

None of the proposals succeeded in developing aligned national standards and tests in any subject, which left states on their own to develop tests. Further there were no mechanisms to help states use the national standards that were developed as a foundation for their own standards. As a consequence, despite well regarded national standards in math, science, civics and other subjects, there was little consistency among state standards, and even greater variations in the quality, rigor and alignment of state tests. Unfortunately, without high quality tests the beneficial impact of good standards is severely weakened.

More coherent federal approaches are readily imaginable. For example, the federal government could develop standards, summative and formative assessments, curriculum and instructional materials, all based on NAEP, and require their use nationwide under NCLB. But the political price for such an approach would be prohibitive. Politically more saleable approaches will of necessity fail to address the coherence challenge.

**The Promise of a State-Led Approach to Common Standards: Lessons From the American Diploma Project**

Notwithstanding federal efforts described above, the standards movement is in fact a state-led effort. Over more than a decade, every state has implemented systems of standards and assessments. More recently states have begun to turn their attention to providing the tools and supports teachers need to implement them in the classroom. The political capital and financial resources that went into developing and building support for these initiatives are significant. Any effort to create national standards must help states build on these investments. They must also help states integrate new or revised standards into a coherent instructional system that aligns professional development, curriculum, formative assessments and data systems and other tools to promote continuous instructional improvement.

States are far better positioned to lead this effort right now. Work underway through Achieve’s American Diploma Project Network demonstrates the potential for collective state leadership to get the job done, faster and more effectively than a federally led effort.

The American Diploma Project—a partnership of Achieve, the Education Trust and the Thomas B. Fordham Foundation—conducted a two year research study with postsecondary institutions and employers in five states to define the knowledge and skills in math and English language arts that high school graduates must have in order to succeed in postsecondary education, and to have access to jobs that pay well and provide opportunities for upward mobility. Based on this study we produced a detailed and specific set of “college- and work-ready” benchmark expectations for the end of high school. We found that “ready for college” and “ready for work” mean the same thing, at least with respect to a core set of academic skills in mathematics, reading and communications that are essential for success in both worlds.

Achieve conducted additional research to
compare state standards with the ADP benchmarks. We found that no more than a handful of states measured up. In fact, we found that no state had a system of standards, curriculum, graduation requirements and accountability that were well aligned with the demands of post-secondary education and work. This “expectations gap”—between the knowledge and skills high school graduates need and what is required of them to earn a high school diploma—is one of the reasons that 30% of first year college students require remediation, and why surveys of college faculty, employers and recent high school graduates themselves reveal that approximately 40% of recent high school graduates are unprepared for either college or work.

At the 2005 National Education Summit on High Schools, convened by Achieve and the National Governors Association, governors from thirteen states, together with state K-12, postsecondary and business leaders, launched the American Diploma Project Network. The purpose of the network is to help states align high school standards, curriculum, assessment and accountability with the knowledge and skills necessary for success in postsecondary education and work. Since then the Network has more than doubled in size; today there are twenty-nine states in the ADP Network that collectively educate more than half of the public school students in the nation.

Achieve is helping teams of postsecondary faculty, employers and K-12 educators from more than 20 states define “college- and work-ready” standards. They are examining national research, state data, employer needs and the work actually done in college classrooms. They are drawing heavily on the ADP benchmarks to begin with, are sharing their work along the way, and have committed to participate in a systematic external quality review of their standards including a comparison with the ADP benchmarks. This won’t guarantee consistency among the states, but the odds are high that there will be quite a bit. As important, this process guarantees a public and transparent external check on state standards, so state officials and educators will know how their own state standards related to the ADP benchmarks.

This is a much different process than what occurred in the 1990s, when virtually every state “based” their standards on national standards, with no independent verification.

The next 10–12 months will tell how much consistency there is among the resulting state standards, though the work to date suggests that there will be a very high degree of consistency among them. The initial ADP research in five states showed what common sense already tells us—the knowledge and skills needed for success in the real world after high school are not dependent upon geography or the boundaries of government jurisdictions. Six additional states have already completed their work, and reviews by Achieve have found that the standards in each state are highly aligned with the ADP benchmarks, and therefore highly consistent with each other.

A number of states in the ADP Network have gone beyond consistent standards, to a common assessment. Nine states have spent the last year working together to develop a common end-of-course test in Algebra II. They recognized that as they increased course-taking requirements for high school graduation, end-of-course exams would be needed to ensure consistent content and rigor across Algebra II courses within each state. By developing a common test the states can create a better test at a lower cost, and compare the performance of their students to those in other states. By aligning this test to the ADP math benchmark expectations, they are creating the pressure to keep the standards and test rigorous, despite the forces in each state, including those generated by NCLB, that tend to exert downward pressures on standards. Finally, this test will be sufficiently rigorous so that postsecondary institutions could use it to let high school students know if they are ready to do college level work, or need to make up skills deficiencies before enrolling. This will help create a more seamless transition to post-secondary education, and help reduce postsecondary remediation rates.
The ADP Algebra II test will be ready for initial administration in Spring 2008. The procurement is designed in such a manner so that additional states can purchase the same test, now or in subsequent years.

Several factors have contributed to the progress these states have been able to make toward consistent state standards and common assessments. These ought to a part of any broader state-led strategy:

The effort is anchored in research on knowledge and skills essential for success in postsecondary education and 21st century work. Anchoring standards in the real world demands students will face after high school has brought an important degree of focus, rigor and consistency to the process. Focusing on the knowledge and skills students need to succeed in college and careers is a powerful source of upward pressure on state standards. The fact that employers and colleges operate in a national and global environment serves to create a high level of consistency across states. While there is still room for considerable collective judgment, the debates about standards in many states are different from debates many of us have grown accustomed to over the years. Conflicts among warring ideological camps in mathematics and English have largely been replaced by evidence-based debates about the advanced math skills really needed to succeed in credit-bearing general education courses compared with those needed to prepare students for the courses that will lead to science, technology, engineering, or math careers. Debates about the literature high school students should read take a back seat to ensuring that students also read complex informational texts and learn to write clear expository writing needed everyday in the workplace.

There must be an external quality review mechanism. The use of external benchmarks and an external review of state standards reinforces the features highlighted above. Further, when differences between state standards and the ADP benchmarks are identified, participants are forced to articulate the reasons for their judgment, making the process and results more transparent than many previous efforts.

The standards must be incorporated into critical state policy tools. To be meaningful, the standards must be incorporated into a core set of state policies and tools in both K-12 and higher education. At a minimum, the standards should form the basis of high school graduation requirements and course descriptions. At the postsecondary level, college-ready standards should provide the basis for placing students into credit-bearing or remedial courses. Placement exams should be aligned with these standards, and postsecondary institutions should consider the information provided by high school assessments aligned with the standards, in making placement decisions.

Implications for the Federal Government

The federal government can provide important support to a state-led effort to create common standards. Briefly, these steps include:

- Through competitive grants, provide financial incentives to states to align high school standards, assessment, curriculum and graduation requirements with the demands of postsecondary education and work, and to align standards and tests in grades 3-8 with these high school standards. To provide adequate phase-in time and to allow states to take different approaches, do not tie new tests to AYP. This work must be done jointly between the K-12 and postsecondary systems, with substantial involvement of the business community; K-12 cannot get this work done working alone.
- Provide financial and technical support to states to work together to develop and use common assessments aligned with common standards.
- Provide financial and technical support to improve the quality and rigor of state tests, and ensure that states are not locked into using existing tests by AYP requirements if they develop more rigorous assessments.
1 Differences in the quality and rigor of state tests are substantial. These differences probably account for most of the variation among states in the percent of students scoring proficient on state NAEP vs. the percent scoring proficient on state tests.

2 States that joined the ADP Network committed to improve their own standards; they were not asked and did not commit to develop or adopt national standards. However, as their work has progressed many have come to see potential for and value of using the ADP Network to create a high degree of consistency in state standards.
No Child Left Behind: A Five Year Review

CONFERENC EPARTICIPANTS

San Juan, Puerto Rico
February 20-25, 2007

Members of Congress

Representative Susan Davis and Steve Davis
Representative Diana DeGette
Representative Bob Etheridge and Faye Etheridge
Representative Sam Farr and Shary Farr
Representative Gene Green and Helen Green
Representative Rush Holt and Margaret Lancefield
Representative Nita Lowey and Stephen Lowey
Senator Richard Lugar and Charlene Lugar
Representative George Miller and Cynthia Miller
Representative Ed Pastor and Verma Pastor
Representative Donald Payne
Representative Janice Schakowsky and Mary Hart
Representative Pete Stark and Deborah Stark
Representative John Tierney and Patrice Tierney

Representative Zach Wamp and Coty Wamp
Representative Melvin Watt and Eulada Watt
Representative Lynn Woolsey

Scholars/Experts

Barbara Byrd-Bennett (Former CEO) Cleveland Public Schools
Michele Cahill Carnegie Corporation of New York
Victoria Chou University of Illinois-Chicago
Michael Cohen Achieve, Inc.
Daniel Fallon Carnegie Corporation of New York
William Galston The Brookings Institution
Richard Ingersoll University of Pennsylvania
Martha Thurlow University of Minnesota
James Wyckoff University at Albany, SUNY
Observers

Alison Bernstein
The Ford Foundation

Paul Goren
The Spencer Foundation

Susan Robinson King
Carnegie Corporation of New York

Adele Simmons
Congressional Program Advisory Committee

Moderator

Dick Clark
Director, Congressional Program
The Aspen Institute

Rapporteur

Carol Copple
National Association for the Education of Young Children

Aspen Institute Staff

Diane Anello
Sara Kuhn
Bill Nell
Pat Walton
No Child Left Behind: A Five Year Review

CONFERENCE AGENDA

San Juan, Puerto Rico
February 20-25, 2007

Education as the Cornerstone of Global Competitiveness: How Does the U.S. Compare Internationally?
William Galston, The Brookings Institution

American competitiveness has become a major area of concern in an age of global trade. How does the education attainment of U.S. students compare with that of other nations, and what trends can be discerned? What is the role of education in maintaining or enhancing American competitiveness? How can students be better educated to meet both the needs of tomorrow's workforce and the challenge of rising education levels around the world?

Adequate Yearly Progress (AYP): Are Absolute Standards or Growth Models the Best Yardstick for Measuring Progress?
Barbara Byrd-Bennett, former CEO, Cleveland Public Schools
James Wyckoff, University at Albany, State University of New York

How do schools currently measure AYP? What are the issues associated with these “status models”? How would “growth models” differ? What approaches will most effectively inform the public and educators about schools’ strengths and weaknesses, create effective incentives for improvement and assist in targeting scarce resources? What types of longitudinal data systems would states and districts need to implement a growth model? What would be required to create such data systems?

Highly Qualified Teachers: How Can We Maximize Their Exposure to Students?
Victoria Chou, University of Illinois at Chicago
Richard Ingersoll, University of Pennsylvania

The right to have a highly qualified teacher was a cornerstone of NCLB. How well are states doing in meeting this requirement, and what are the major obstacles in doing so? Is the teacher shortage one of the major problems? What do the data tell us about the teacher shortage, the reasons behind it and the solutions to it? Is the current definition the most effective approach to promoting teacher
quality? What is the correlation between strong qualifications and being a good teacher? Are there ways to measure teacher effectiveness? What policies—concerning compensation, working conditions, certification, teacher assignment and transfer provisions, etc.—can help districts recruit and retain effective teachers in hard-to-staff schools and subjects?

The Challenge of Special Populations to Accountability for All
Martha Thurlow, University of Minnesota

Students with disabilities and English learners are challenging the vision of accountability for all and the goal of leaving no child behind. Yet, child advocates and others claim that NCLB is important for the educational success of these children. What does the evidence say about their student outcomes under NCLB? What factors related to educators’ beliefs, school capacity, and assessment design play into the challenges of special populations achieving accountability for all? Where do the solutions lie? Should special populations face the same standards and assessments, or do they merit exemptions?

Does Education Require National Standards?
Michael Cohen, Achieve, Inc.

Discrepancies between state and federal test results lead some to conclude there is a need for national standards. What are the pros and cons of this? Is it politically feasible? Are there alternative approaches that will help ensure that all state standards are sufficiently rigorous that students graduate prepared for college and work?