

Value-Added Assessment and Systemic Reform: A Response to the Challenge of Human Capital Development

Mr. Hershberg believes that the highly competitive global economy of the 21st century requires dramatic improvements in America's schools. While he sees merit in the basic intent of No Child Left Behind, he suggests that a shift to value-added assessment, if it is used as the foundation of a comprehensive school reform model, will be the key to raising the achievement of all students.

BY THEODORE HERSHBERG

Our nation, which has prevailed in conflict after conflict over several centuries, now faces a stark and sudden choice: adapt or perish. I'm not referring to the war against terrorism but to a war of skills — one that America is at a risk of losing to India, China, and other emerging economies. And we're not at risk of losing it on factory floors or lab benches. It's happening every day, all across the country, in our public schools. Unless we transform those schools and do it now . . . it will soon be too late.

— *Louis Gerstner*, former chairman, IBM; chairman, The Teaching Commission

TERRORISM and the war in Iraq are high on the list of the nation's concerns, but the greatest danger facing America is, as Louis Gerstner recognized, the challenge of human capital development. Our nation's public schools, the foundation for this effort, are still failing far too many of our children despite an investment of some \$500 billion annually.

Sadly, we've known about this threat for quite some time. In 1983 *A Nation at Risk* suggested that, if the mediocrity of our schools had been imposed by an unfriendly foreign power, we might well have viewed it as "an act of war." We understood then that companies all over the world could buy foolproof machinery that compen-

THEODORE HERSHBERG is a professor of Public Policy and History and executive director of Operation Public Education, University of Pennsylvania, Philadelphia. He was appointed to the national working group to advise U.S. Secretary of Education Margaret Spellings on incorporating student growth data into the implementation of the No Child Left Behind Act. This article was originally prepared for the Aspen Institute's Congressional Institute, "The Challenge of Education Reform: Standards, Accountability, Resources, and Policy," Cancun, Mexico, February 2005.

sated for deficient worker skills and that billions of people were willing to use that machinery for a fraction of the wages American workers wanted. But recent trends make clear that globalization is about far more than displacing our nation's blue-collar workers.

Every job in America is at risk. As author and *New York Times* columnist Thomas Friedman argues in *The World Is Flat*, 30 years ago you'd rather be a B student in Boston than a genius in Bangalore or Beijing because the economic opportunities were all in America. Today, he explains, that is no longer true. Intellectual work can be digitized, disaggregated, delivered, distributed, produced, and shipped from anywhere on the planet to anywhere else.¹

Friedman identifies 10 developments — or “flatteners” — that have helped to level the playing field of the global economy. The fall of the Berlin Wall in 1989 signaled the triumph of capitalism. The development of Microsoft Windows and the introduction of the Netscape browser helped fuel the high-tech bubble. Most of us looked at the dot.com collapse as a disaster, but it was actually a great boon to globalization because billions of dollars were spent on fiber optic telecommunications cable to connect the continents. When the initial investors went bust, other companies bought these transmission networks for pennies on the dollar, which made it possible to offer those networks to users worldwide at minimal cost. New software created compatibility among diverse computer applications. Companies could now send jobs (outsourcing) or whole factories (off-shoring) overseas; have employees in different locations collaborate online through shared operating systems, such as Linux (open sourcing); allow other companies such as UPS to take over whole sections of their operations (insourcing); or create global supply chains so that when a product is sold in the U.S., another is immediately made in China (supply-chaining). Powerful new search engines, such as Google and Yahoo, enabled people around the world to mine unlimited data sources, and new wireless technologies and Voice Over Internet Protocols are now accelerating these developments.

The 21st century has seen three billion new people enter the world economy, the majority from China, India, and the former Soviet Union, all societies with rich educational heritages. China and India alone each year graduate 5.1 million students from college, including 400,000 engineers, while our schools turn out 2.2 million college graduates and 60,000 engineers. We produce only two bachelor's degrees for every 10 students who start high school. “The law of sheer numbers,” observes Hewlett-Packard's former CEO Carleton Fiorina, “is fairly compelling.”

The dirty little secret that no CEO wants to say publicly is that they are already outsourcing — not just because

they can save money on salaries, but because they can often find better-skilled, more productive, and more ambitious people overseas than in America. Bill Gates is investing millions of dollars to improve our nation's high schools because he is “terrified” for the future work force of the nation. “In the international competition to have the biggest and best supply of knowledge workers,” Gates declares, “America is falling behind.”

The consequences for failing to meet the human capital development challenge are already severe. The last year the typical blue-collar worker earned enough for mom to stay at home and raise the kids was 1964. We maintained our standard of living, despite the decline of real wages through these decades, largely because women entered the labor force in record numbers and created two-income families. But that strategy has run its course. Virtually all gains in family income in the last 30 years have gone to the top one-fifth of American families, largely as the result of new technologies that favor the better educated.

NO CHILD LEFT BEHIND

Our last four presidents, the Congress, governors, and corporate leaders have come to understand that, if America is to remain a stable, middle-class society, steps must be taken to significantly improve our system of public schools. Frustrated by precious little improvement in student achievement over the last two decades, Congress in a bipartisan consensus passed the No Child Left Behind (NCLB) Act in 2001. NCLB was an unprecedented expansion of the role of the federal government in K-12 education.

NCLB broke new ground. It required states to set academic standards so that we could move away from norm-based testing, which compared students to one another, and learn instead what students know and can do at given ages. It required states to test annually so we could have a basis for measuring changes in achievement levels. It forced schools to focus attention on the academic progress of long-neglected low-income and minority students, which is why many believe the law is one of the great civil rights victories in our history. And it introduced consequences for schools that failed.

The legislation was necessary, and, despite some design flaws that will be addressed later, we must not retreat from its goals. Nonetheless, it would be seriously wrong to conclude that NCLB represents a sufficient response to the human capital development threat posed by globalization.

NCLB was not and — given political realities at the federal level — could not in itself be a dramatic overhaul of the K-12 system. We remain the only developed nation in

which the national government neither dictates nor closely coordinates the standards, assessments, and curricula that are used in its schools. Yet the reason for breaking the long-standing “hands-off” role of the federal government was that, to meet the challenge of human capital development and to ensure the civil rights of children, our schools had to do something never done before: educate *all* children, not just the top one-fifth, and educate them to unprecedentedly high levels (or levels never before imagined).

NCLB cannot directly produce fundamental change in American public education, but it can help the states transform our nation’s public schools. What follows is a description of the kind of system needed and a discussion of how we can work with NCLB to encourage the states, through incentives and regulatory modifications, to promote the required changes so that schools can graduate students who are able to use technology, think critically, solve problems, and learn throughout their lifetimes. Whether this can really be accomplished by each state remains to be seen, but at this moment, barring a sea change in the Congress, that is the best we can hope for.

AMERICA’S PUBLIC SCHOOLS: PAST AND FUTURE

The system of public education now in place is largely unchanged from the 19th century, when schools were set up to do three things. The first was to provide universal basic literacy, and America became the first nation in which virtually everyone in the labor force could read and write at the sixth-grade level. The second was to socialize a highly diverse population — millions of immigrants from different nations, cultures, and religions and millions of farmers who migrated to cities — for success in an industrial economy. Students were taught to show up on time, respect authority, develop a work ethic, and repeat monotonous tasks. Third, using standardized tests and the bell-shaped curve, the schools identified and sorted out the top one-fifth of their students for higher education, and the best and brightest of these went on to run the country.

The old system, concerned with quantity and cohorts rather than quality and individuals, was designed to let the cream rise to the top. For the remaining 80% or so, there was little consequence because, for most of our history, they ventured forth into an industrial economy that provided ubiquitous jobs paying middle-class-sustaining wages but requiring little in the way of advanced education or higher-level skills.

Our schools were enormously successful in these tasks, and there is no way to understand the emergence of America as an industrial superpower without acknowledging

the key role they played. The problem is that people continue to behave as if the current school system — designed for a different century and a different economy — is the right one to meet the challenges ahead, despite the record of the last three decades. Since 1970, notwithstanding an increase in real spending of over 100% per pupil, a decrease of 22% in the pupil-to-teacher ratio, and a doubling of the number of teachers with master’s degrees, student achievement has remained largely flat.

For schools to succeed in the 21st century, teachers and administrators must undergo the requisite education and training to master concepts and methods that were not needed in the schools of an industrial economy. Educators will need to replace the “ability-based” notion that dominates the thinking of too many who work in our schools with an “effort-based” theory of learning: you are not simply born smart, but you can “get smart” with appropriate effort, resources, and high-quality instruction. They will need to see how teaching a standards-driven curriculum is very different from using a bell-shaped curve to distinguish among students. Today, all students must reach high standards. Teachers will need to be proficient in using a problem-solving pedagogy in the classroom: memorization will always have a place in the learning process because no subject can be mastered without it, but it must never again be the dominant feature of schooling. Teachers will have to learn how to “differentiate instruction”: one size definitely does not fit all in an era in which we can afford to leave no child behind and children have many different styles of learning. They must master data-driven decision making, a striking departure from the anecdotal approach that has long characterized our schools. Finally, teachers and administrators must learn how to shift from teacher-centered to student-centered classrooms: lectures serve the needs of some students, but we all learn best when we bear more of the responsibility for learning.

But before our educators can master these new skills and knowledge, key problems that plague the profession of teaching must be addressed. According to recent reports from the two national commissions on teaching, one in three teachers leaves in the first three years; 46%, in the first five years. Moreover, these rates are 50% higher in urban districts, and the qualifications profile of the “leavers” is stronger than that of the “stayers.” Of the 3.4 million teachers today, two million will leave in the next decade — three times as many through attrition as through retirement. To replace these teachers and to retain and attract the “best and the brightest” will require a transformed school system, one that will make teaching a more financially rewarding and intellectually satisfying experience.

A NEW SYSTEM: OPERATION PUBLIC EDUCATION'S COMPREHENSIVE REFORMS

The system necessary to encourage and support these changes will have to be governed through an entirely new set of rules and incentives. Many key elements of reform — addressing assessment, educator quality, compensation, professional development, and capacity building — are already in place across the country. The challenge is to bring all of these together, along with several striking innovations, so that the typical practice of piecemeal reform that produces marginal improvements is replaced by an effort that can yield sustainable systemic change.

Operation Public Education (OPE), based at the University of Pennsylvania, has done just that. It has developed new and comprehensive reforms that complement the federal law and provide tools that can transform America's schools. These reforms have been codified into legislative language and are now being promoted nationally with support from the Annenberg Foundation and the Carnegie Corporation. Winning support for these progressive and controversial reforms is a difficult task, but it is not impossible. Districts in several states have expressed a genuine interest in piloting our reform model.

The OPE system provides new forms of educator evaluation that include *outputs* (student learning results) in addition to *inputs* (the observation of teachers in their classrooms). The compensation system enables outstanding teachers to earn higher salaries more quickly and is flexible enough to differentiate pay for those difficult-to-fill vacancies associated with particular subjects or less desirable working environments. It provides more fluid career opportunities so that effective teachers can assume greater responsibilities at earlier ages. Teachers who need remediation are required to undergo it, and ineffective teachers who are unable to improve must leave the profession. Professional development opportunities are substantially expanded so that educators can continue to grow throughout their careers.

At the heart of the OPE system is an essential *quid pro quo* in which teachers accept accountability as individuals in return for a significantly expanded role in school management: 1) teachers are given responsibility for evaluating fellow educators through a peer-review process; 2) they play a key role in the remediation process that affects all key personnel decisions; and 3) because their status is now determined by what happens in their classrooms rather than at the bargaining table, they become full partners in the policy decisions that affect classrooms, such as those concerning professional development and curricula. While collective bargaining remains in place in the OPE system,

the hard distinction between teachers as “labor” and administrators as “management” is dissolved because the difficulty of systemic reform requires close cooperation of all key parties.

Finally, the mediocre, high-stakes standardized tests found in the large majority of states need to be replaced with a new integrated assessment system that would provide not only a high-quality “summative” exam at year’s end, focused on the development of higher-order thinking skills, but also “formative” assessments throughout the school year that are designed to give teachers regular feedback in the form of suggested pedagogical interventions to support improved instruction for this year’s students. OPE is working with the National Center for the Improvement of Educational Assessment as well as with representatives of all the major test-publishing companies to develop a request for proposals to be issued by states and large school districts calling for the creation of just such an integrated testing regime.

VALUE-ADDED ASSESSMENT AND ACCOUNTABILITY FOR INDIVIDUAL EDUCATORS

The OPE system rests on a foundation of individual-level accountability — both teachers and administrators are held responsible as individuals for student learning results. NCLB moved in the right direction in requiring accountability. But in making the school rather than the individual educator the unit of accountability, it fell short for two important reasons. First, because there is greater variation in the quality of instruction *within* schools than *between* them, it is essential to report data at the classroom level for evaluation purposes. Second, systemic changes of the type discussed above will be achieved only when the careers of everyone working in our public schools are tied to successful learning outcomes.

But until now, efforts to hold individual educators directly responsible for student learning — sometimes referred to as “pay for performance” or “merit pay” — have failed to differentiate effectiveness either because they were too subjective or because they relied on achievement test scores, which are strongly influenced by family circumstances (income level, types of jobs, years of schooling, attitudes about education, and so on). The unfairness of such approaches was recognized, and they have largely been abandoned.

At a recent national conference, Dan Fallon, chair of the Education Division of the Carnegie Corporation, explained the origins of the belief held by most Americans — laypeople as well as K-12 educators — that the level of academic achievement is determined largely by factors

beyond a school's control. It can be traced back to James Coleman's 1966 report, in which Coleman found that "only a small part of [student achievement] is the result of school factors, in contrast to family background differences between communities," and to the work of Christopher Jencks in 1972, in which he concluded that "the character of a school's output depends largely on a single input, namely the characteristics of the entering children."² This understanding is reinforced for the public at large when metropolitan newspapers issue their annual "report cards" on the schools, which reveal that wealthy communities almost always have the highest test scores.

"An implicit conclusion of the analyses put forward by Coleman and Jencks," Fallon observed, is that "when it comes to student achievement, teaching doesn't matter very much."³

But a spate of new studies now proves empirically that teaching matters enormously. To understand how this new conclusion can be arrived at, it is vital to grasp a fundamental distinction within the measurement of student learning. *Achievement* refers to the absolute levels attained by students on their end-of-year tests. *Growth*, in contrast, refers to the progress in test scores made over the course of the school year. And here is the most important implication of this difference: high absolute scores on assessments such as the SAT are best predicted by family income, but if we are predicting student growth — progress made over the year — reports by education researchers John Kain, Eric Hanushek, William Sanders, and others have demonstrated that good instruction is 15 to 20 times more powerful than family background and income, race, gender, and other explanatory variables.

Given the technological limitations of their era, Coleman and Jencks focused on what they could: achievement. But today, because researchers have access to datasets and technology that can link the progress of individual students over time to the teachers who taught them, it is now possible to measure the impact of instruction on a student's academic growth by using a powerful new methodology called *value-added assessment*.

Value-added assessment is often confused with simple growth because the words themselves make it is easy to think about this growth as the "value" that is "added" over the last year. But the statistical method known as "value-added assessment," as developed for the state of Tennessee by William Sanders when he was a professor of statistics at the state university, is a way of isolating the impact of instruction on student learning. Its great advantage is its ability to separate the annual academic growth of students into two parts: that which can be attributed to the student and that which can be attributed to the classroom, school,

or district. Because individual students rather than cohorts are tracked over time, each student serves as his or her own "baseline" or control, which removes virtually all of the influence of the unvarying characteristics of the student, such as race or socioeconomic factors.

Although there are several different value-added models in use today, only the Sanders model has been mandated for use statewide: in Tennessee since 1992 and, most recently, in Pennsylvania and Ohio. It is used as well in over 300 other school districts in 21 states.

Under the value-added approach, test scores are projected for students and then compared to the scores they actually achieve at the end of the school year. Classroom scores that exceed projected values indicate effective instruction. Conversely, scores that are mostly below projections suggest that the instruction was ineffective.

At the same time, this approach considers student factors, such as the pattern of prior test scores, both those of the individual student and those of other students in the same class. If a student's present performance is below projected scores, while students with comparable previous academic histories in the same classes have done well, this is evidence of the student effect — that is, external variables such as the home environment — which is outside the control of teachers and schools.

Since students' projected scores are based only on their prior academic records rather than on race or socioeconomic background, value-added assessment does not introduce bias. In other words, low-income children are not expected to do poorly and high-income students are not expected to do well. But because value-added assessment tracks the same students over time — thus accounting for family and neighborhood characteristics that so strongly bias absolute test scores — educators are not penalized for circumstances beyond their control.

When value-added scores are collected for each classroom and averaged over three years, teachers have rich diagnostic information to improve their instruction, and administrators have an empirical basis for evaluating teacher effectiveness. When these classroom scores are aggregated over entire buildings and districts, principals and superintendents can be held accountable for student-learning results.

Value-added assessment is not without controversy, given the complicated statistics on which it unavoidably rests. The RAND Corporation's major study of value-added models identified important research questions about school effects, about the comparability of the instructional difficulty at different grade levels and in different subjects, and about the quality of the tests used. Nonetheless, it conclud-

ed that 1) the teacher effect is real, 2) it could be large, and 3) it persists beyond the year in which it is first evident. The RAND researchers concluded that value-added models (VAM) “might actually provide less-biased and more precise assessments of teacher effects” than existing test-based systems and that as policy makers “evaluate alternative models for school or teacher accountability, VAM should be given serious consideration even in light of its limitations.”⁴

In a balanced accountability system, no educator should ever be evaluated solely on the basis of a single measure, not even one as powerful as value-added models provide. In the OPE system, for example, value-added scores constitute half of an educator’s evaluation, while the other half is based on direct observation of performance, using the well-established protocols developed by Charlotte Danielson.⁵ Value-added models used for accountability should always be accompanied by safeguards, such as those developed by OPE, to ensure that educators are treated fairly as individuals. Given the nation’s human capital development challenge, we must be willing to innovate, take risks, and not let the perfect be the enemy of the good.

SOME PROBLEMS WITH ADEQUATE YEARLY PROGRESS

NCLB requires schools to bring all children up to high standards by 2014. That is certainly a worthy goal. So, too, is the insistence that raising schoolwide averages is not enough: student subgroups — including low-income, non-English-speaking, and special-needs students, as well as those of varied ethnicities — must meet these standards as well. The problem, however, is how to identify which schools are on target to meet these requirements. In most cases, NCLB’s adequate yearly progress (AYP) measures can correctly sort out successful schools from those that are failing their students. But for many schools, AYP measures do not provide a fair and complete assessment of school performance.

At the heart of this problem is the fact that AYP focuses on *achievement* to the exclusion of *growth*. Table 1 helps us identify and understand the twin deficiencies of AYP.

Proficiency (i.e., achievement) is tracked on the vertical axis, while *growth* is tracked on the horizontal axis. In the bottom left cell are schools that are clearly not serving the needs of their students — they are producing both low proficiency and low growth. Clearly, these schools deserve to be sanctioned. Schools in the top right cell are performing wonderfully. They are doing what we want all schools to do: providing their students with both high proficiency

and high growth. For the schools in these two cells, simple AYP measures accurately reflect their educational outcomes.

Unfortunately, not all schools fall in these two cells. The problems with AYP are clearly evident in the remaining two cells. In the top left cell are schools whose students are meeting their AYP goals, but little growth is occurring. Most such schools are found in affluent communities, where high test scores go hand in hand with high family income. These schools can be referred to as “slide-and-glide” schools because they rest easily on the laurels of their students. It is important to understand that NCLB does nothing to hold these schools accountable for providing their students with the annual growth to which they are entitled. In a global economy characterized by fierce competition for demanding jobs that pay high salaries and benefits, this is a highly significant shortcoming.

The bottom right cell contains schools with high growth but low proficiency. These schools have clearly succeeded in “stretching” their students academically, but, given how far behind the students were when they entered school, the schools have not yet been able to raise them to AYP-required levels. Thus they still face sanctions under current law.

When NCLB was drafted, some educators advocated the inclusion of growth in the measurement of AYP so that schools that helped their students grow would be seen as doing a good job, but their proposals were turned down. This decision has fed into a conspiracy theory widespread in the K-12 world that sees NCLB as a plot to make the nation’s schools look bad (by, for example, ensuring a high rate of failure to make AYP) so citizens will give up on public education, thus opening the door for vouchers.

The truth, of course, is very different. The refusal to include growth as an alternative for meeting AYP was a bipartisan decision. Supporters believed that, even if students — primarily of minority and low-income backgrounds — grew every year, many would never reach real-world stan-

**TABLE 1.
Possible School Outcomes**

| | | |
|-----------------------|--------------------------------|---------------------------------|
| ↑ Proficiency ↓ | High proficiency Low growth | High proficiency High growth |
| | Low proficiency Low growth | Low proficiency High growth |
| | ← Growth → | |

dards by graduation because they had entered school so far behind. The position of NCLB's backers was quite principled — they were not going to “abandon these children.” So they rejected growth measures and insisted that the only way schools could make their AYP targets was to get all their students to proficiency.

CAN ALL CHILDREN ACHIEVE HIGH STANDARDS?

Some observers believe that most children who enter school well below grade level will be unable to reach proficiency by graduation. In *Class and Schools*, for example, Richard Rothstein argues that factors beyond the control of schools have such an overwhelming impact on student achievement that only a massive infusion of new funds can overcome the major deficiencies in health, nutrition, socialization, and income that hobble these students.⁶

Others are agnostic: maybe all students can, maybe all students can't reach high standards. For now, they argue, schools should focus on ensuring that all students get a year's worth of growth every year from wherever they start in September. Knowing that schools fall far short of this goal each year — for example, 68% of K-8 schools in Tennessee did not provide students with a year's worth of growth in math in 1996-97⁷ — they conclude that, if schools simply did not lose ground, students would be vastly better off.

Of course, sustaining cumulative gains over years of schooling would be a dramatic improvement over the status quo. A student entering second grade on grade level in a school averaging 75% annual growth would graduate from eighth grade at 6.5 years of academic achievement, while one attending a school averaging 140% would graduate at 10.4 years of academic achievement.

Still others, building on this idea, argue that an achievable and worthy goal for the nation's schools would be to require them to “stretch” their students beyond a year's worth of growth in a year. That is, not only would the students not lose ground, but they would grow at a rate that exceeded predicted performance based on earlier achievement. This accomplishment, they argue, should qualify schools as meeting AYP, even if the level of achievement should fall short of proficiency.

But because NCLB is the law of the land, this is currently a philosophical discussion. Time will tell whether our schools can, in fact, bring all children to proficiency by graduation. Adopting OPE's comprehensive reform model will increase the likelihood of this outcome. But for now let us consider how it is possible to improve the idea of AYP without abandoning the commitment that all children reach proficiency. This approach is called “growth to stan-

dards,” and it is achievable through the introduction of value-added models.

FIXING AYP WITHOUT ABANDONING PROFICIENCY

The essence of the growth-to-standards approach is identifying schools that are putting their students on growth trajectories that will reach proficiency in the future and crediting these schools for that achievement.

Schools would demonstrate performance by using a value-added methodology that converts the static achievement scores of their students to dynamic growth scores. If students currently performing below their AYP target are on track to reach proficiency by the time they graduate, they would be counted among those meeting their AYP target in the current year. If a school were to place enough of these students on growth-to-standards trajectories, it could meet its AYP goal for the year. Using a growth-to-standards approach, in other words, would reduce the proportion of schools failing to meet AYP, but without abandoning the commitment to proficiency. Several existing approaches — the Northwest Evaluation Association's growth model, Harold Doran's REACH model, and William Sanders' value-added model — could accomplish this goal.

This approach may be criticized for the same reason that the existing definition of AYP is criticized: it creates what many call a “perverse incentive” for educators to focus like a laser beam on one group of students to the exclusion of all others: those close to but below proficiency. Schools could choose to ignore students far below proficiency as well as those whose scores already exceed proficiency, the argument goes, because the prime directive in NCLB is for schools to hit their annual AYP targets.

While this is clearly the logic of the incentive, we do not yet know if it is supported in fact. The growth-to-standards approach described above, like AYP, might simply illuminate the pattern — the gains made by those who start just below proficiency are coming at the expense of those who start the year above it — rather than exacerbate it.

We know this pattern long predates NCLB and has been widespread in poor communities, whether in inner cities or in Appalachia. It explains, for example, the observation made by elementary school teachers that the proportion of precocious students in kindergarten and first grade is sharply reduced by fifth and sixth grades. Faced with so many low-performing children, the explanation goes, teachers focus on the bottom of the student distribution so that previously low achievers show high growth, while previously high achievers show low growth. Sustaining this focus in the early years explains why so few high-achieving,

low-income children are found in middle school.

When William Sanders applied his growth-to-standards approach to all Tennessee schools in the 2002-03 school year, he learned that 13% more schools would meet their federal goals if this alternative means of calculating AYP were accepted by the U.S. Department of Education. But when Sanders looked more closely at its effects — he examined nine Memphis schools, all of whose students were minority and low-income (i.e., receiving free and reduced-price lunch) — he discovered some troubling results. While some schools met their AYP through the growth-to-standards alternative without denying any of their students adequate yearly growth, others did so at the expense of students who had achieved at higher levels in the past. Seeing no sense in a tradeoff that benefits one group of poor minority kids at the expense of another, Sanders has proposed a “net” approach: schools would receive credits for students placed on a growth-to-standards trajectory and debits for formerly higher-achieving students denied adequate growth in the process.

The U.S. Department of Education should encourage states currently using high-quality value-added or growth models to conduct pilot studies over the next few years to determine the precise impact of a growth-to-standards approach.

CONCLUSION

It is essential that school reformers understand the limits of what can be done at the federal level, so they can concentrate their efforts on the states. The essential focus of NCLB was on bringing up the bottom — itself a significant and long-neglected goal of social justice. It was not — and, given political realities, could not have been — about ensuring that the bar was raised sufficiently high so that in the future all students would graduate as well-educated citizens in an increasingly complex society and as productive workers in the highly competitive global economy of the 21st century.

The political compromises necessary to get NCLB passed left the height of the proficiency bar to the states. Comparing the results for students considered proficient in state tests with results from the National Assessment of Educational Progress makes clear that, with a few notable exceptions, state standards were set low to begin with. To meet their goals under the new federal law, some states are watering down their proficiency targets. Others have “back-loaded” their goals, setting the time when the highest percentage gains in student performance are to be made well in the future in the hope that NCLB will no longer exist.

Still others are sitting by passively and allowing schools to subtly — and not so subtly — encourage their lowest-performing students to drop out. And in the face of high failure rates, states are being allowed to use “confidence intervals” as a way of measuring AYP, a statistical adjustment that further waters down their AYP requirements.

Therefore, we need to create incentives at the federal level to encourage the states to move in different and more desirable directions, because it is at the state level that the school system of the 21st century will be created. I’ve argued that value-added assessment could serve as a powerful catalyst for change, and momentum for its use is building. Beyond Tennessee, Pennsylvania, and Ohio, where the Sanders model has been adopted for statewide use, Arkansas, Minnesota, Colorado, and Florida have passed legislation calling for the introduction of value-added models. These states have recognized the many advantages of value-added assessment: it traces the academic progress of individual students, rather than cohorts; it focuses on ensuring that all students, not simply the lowest performers, receive at least a year’s worth of growth in a year; it provides educators with rich diagnostics to improve instruction; and, if used in a growth-to-standards approach — which the federal government is leaning toward embracing — it would help schools legitimately meet their AYP goals without abandoning the new federal commitment that poor and minority children should graduate from school having met proficiency standards.

But the real prize in having value-added assessment widely adopted is that, as the basis for individual-level accountability, it serves as the foundation for comprehensive reforms, such as OPE, that will change the organization and governance of K-12 schools. Only then will we have in place the means to bring all American students to internationally competitive standards — a vital step toward meeting the nation’s human capital development challenge.

1. Thomas L. Friedman, *The World Is Flat: A Brief History of the Twenty-First Century* (New York: Farrar, Straus & Giroux, 2005).

2. Daniel Fallon, “Clarifying How We Think About Teaching and Learning,” paper presented at the Battelle for Kids National Value-Added Conference, Columbus, Ohio, 25-26 October 2004. Available at www.battelleforkids.org.

3. Ibid.

4. Daniel McCaffrey et al., *Evaluating Value-Added Models for Teaching Accountability* (Santa Monica, Calif.: RAND Corporation, 2004), p. 120.

5. Charlotte Danielson, *Enhancing Professional Practice — A Framework for Teaching* (Alexandria, Va.: Association for Supervision and Curriculum Development, 1996).

6. Richard Rothstein, *Class and Schools: Using Social, Economic, and Educational Reform to Close the Achievement Gap* (New York: Economic Policy Institute/Teachers College, 2004).

7. *Graphical Summary of Educational Findings from the Tennessee Value-Added Assessment System* (Knoxville: University of Tennessee Value-Added Research and Assessment Center, 1996). 

File Name and Bibliographic Information

k0512her.pdf

Theodore Hershberg, Value-Added Assessment and Systemic Reform: A Response to the Challenge of Human Capital Development, *Phi Delta Kappan*, Vol. 87, No. 04, December 2005, pp. 276-283.

Copyright Notice

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

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

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