
Clinton Presidential Records Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a publication.

Publications have not been scanned in their entirety for the purpose of digitization. To see the full publication please search online or visit the Clinton Presidential Library's Research Room.

PRINCIPAL

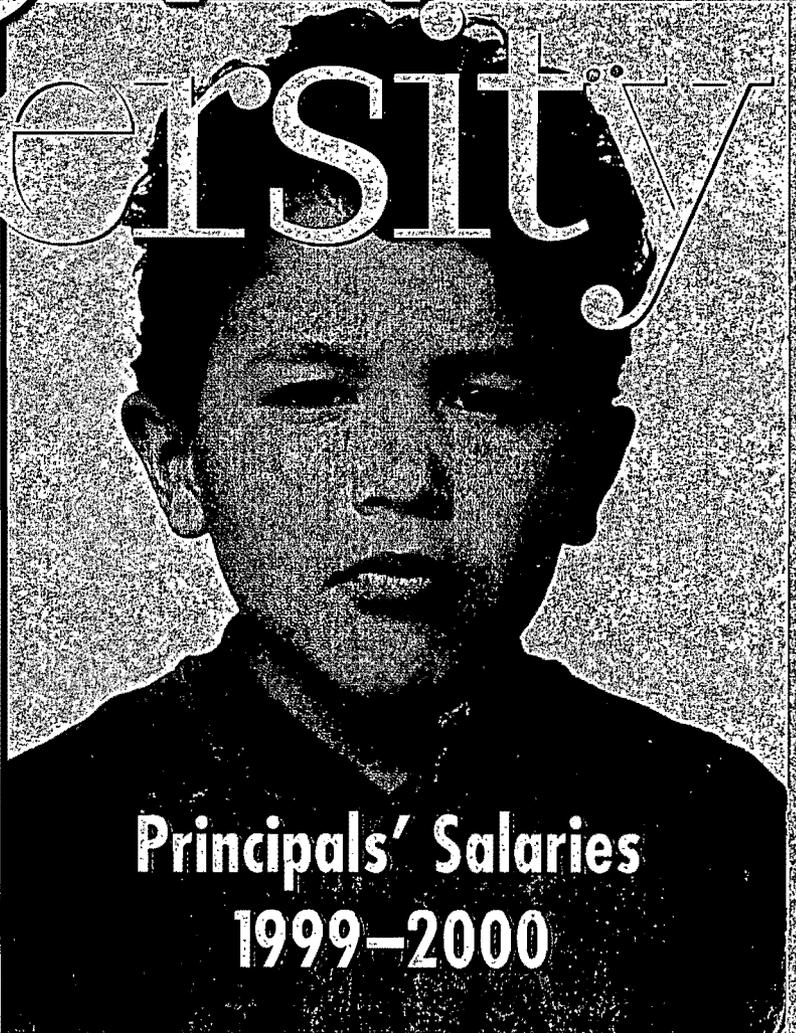
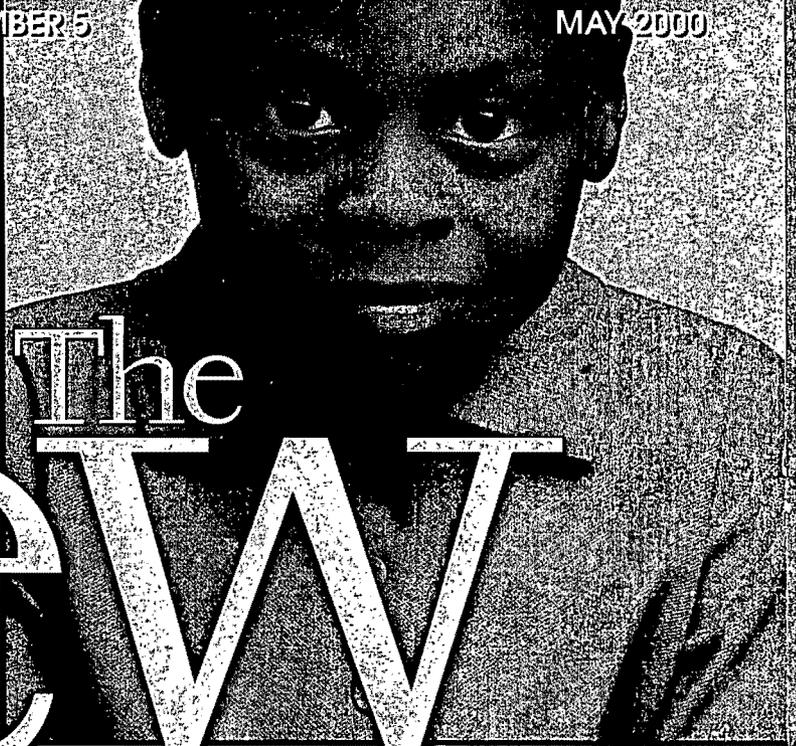
VOLUME 79

NUMBER 5

MAY 2000

THE NEW

Diversity



*****MIXED ADC 050-500049391 P 1785 S 494
OFFICE OF POLICY DEVELOPMENT
OLD EXECUTIVE OFFICE BLDG.
17TH ST. & PENNSYLVANIA AVE, NW
WASHINGTON, DC 20500

Principals' Salaries 1999-2000

Clinton Presidential Records Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a publication.

Publications have not been scanned in their entirety for the purpose of digitization. To see the full publication please search online or visit the Clinton Presidential Library's Research Room.

Clinton Presidential Records Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a publication.

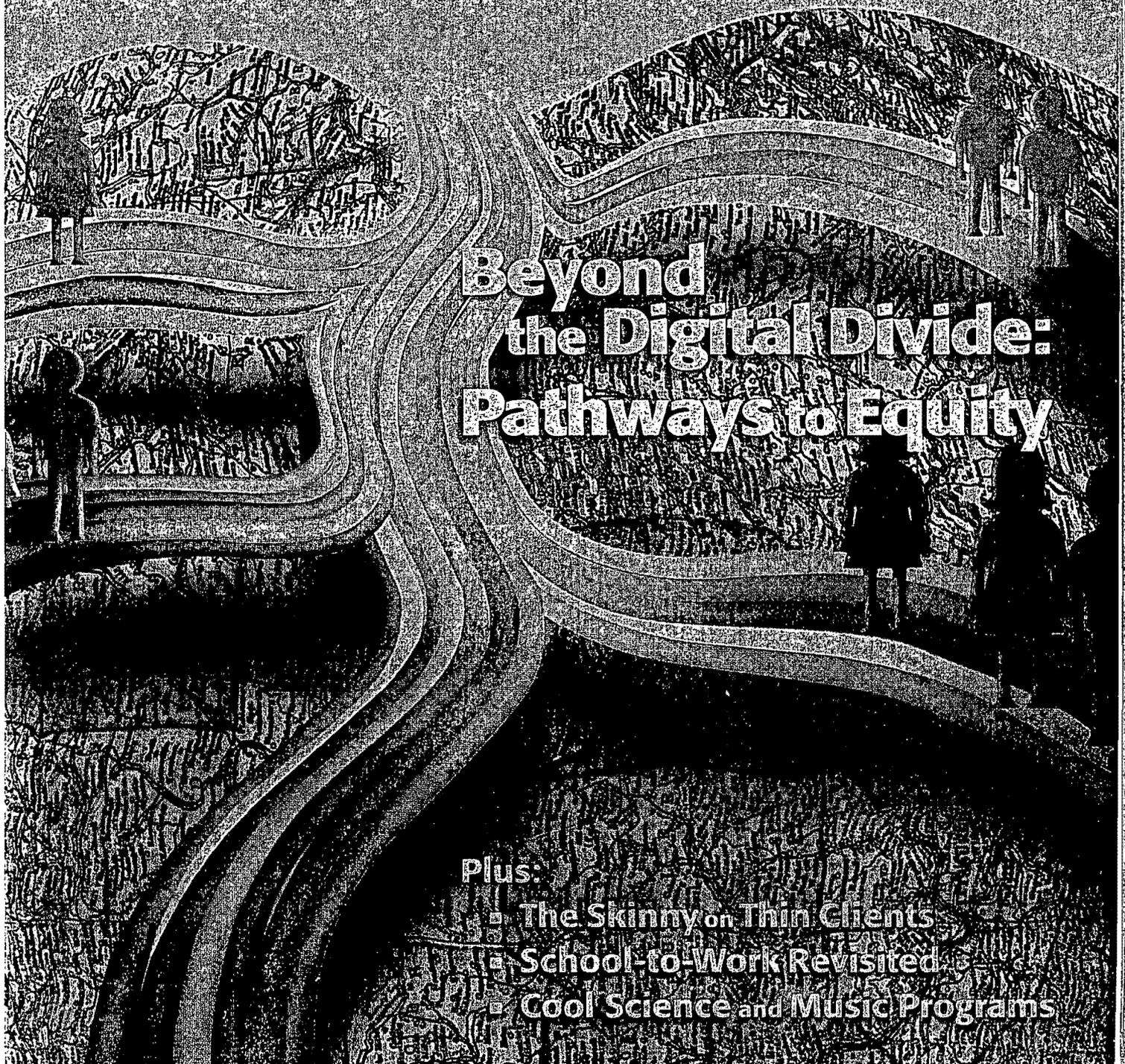
Publications have not been scanned in their entirety for the purpose of digitization. To see the full publication please search online or visit the Clinton Presidential Library's Research Room.

MP

1980-2000 **20th** Anniversary

TECHNOLOGY & LEARNING

MAY 2000 • VOLUME 20 NUMBER 10 • \$4.00



Beyond the Digital Divide: Pathways to Equity

Plus:

- ▣ The Skinny on Thin Clients
- ▣ School-to-Work Revisited
- ▣ Cool Science and Music Programs

Clinton Presidential Records Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a publication.

Publications have not been scanned in their entirety for the purpose of digitization. To see the full publication please search online or visit the Clinton Presidential Library's Research Room.

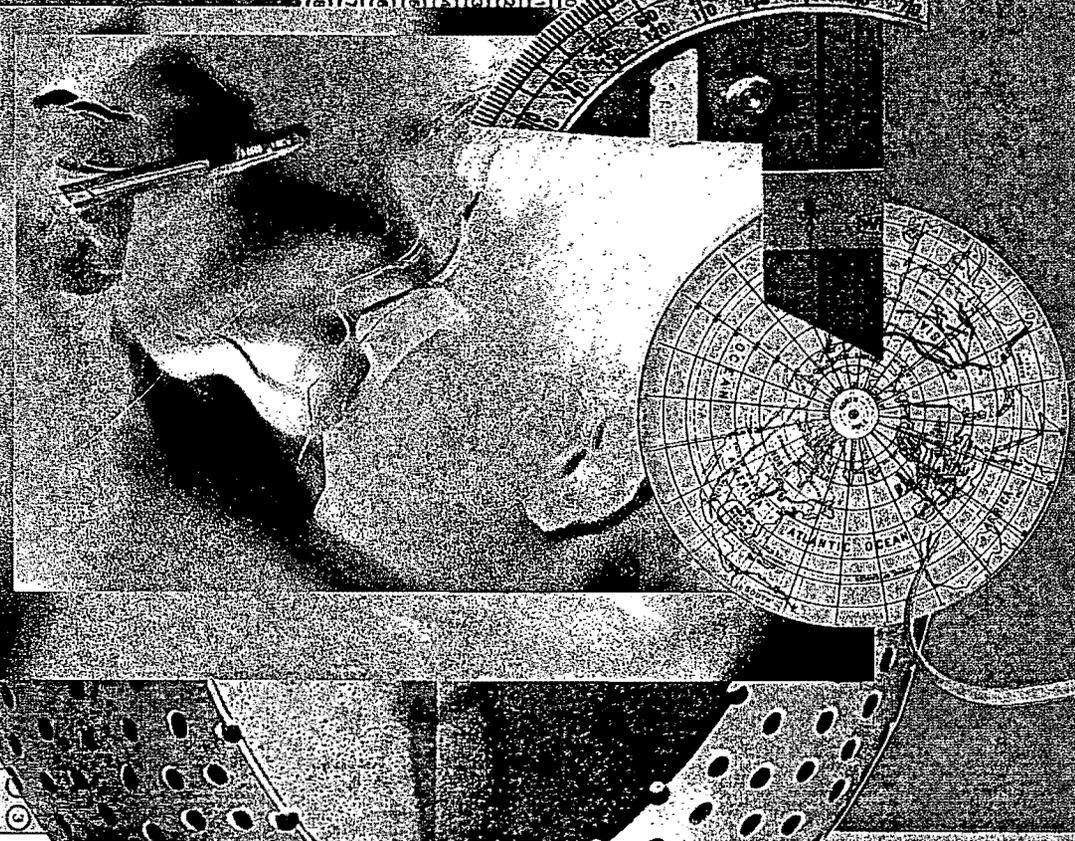
American School Board

JOURNAL

JANUARY 2000

Tests that Count

Do high-stakes assessments
really improve learning?



*****ALL FOR ADC 200
HLIST 429 3361
THE WHITE HOUSE
OFFICE OF POLICY DEVELOPMENT
WASHINGTON DC 20500



Talking About Education

Looking for inspiration in the new century—or a fresh quote to end your e-mail?

Try some of these words from the distant, and not-so-distant, past

Chris Trippi

"A child miseducated is a child lost."

—John F. Kennedy, U.S. President (1917-63)

"Count it one of the highest virtues upon earth to educate faithfully the children of others."

—Martin Luther, German religious leader (1483-1546)



"Many things we need can wait, the child cannot. Now is the time his bones are formed, his mind developed. To him we cannot say tomorrow, his name is today."

—Gabriela Mistral, Chilean writer (1889-1957)

"Give me the children until they are seven and anyone may have them afterwards."

—Francis Xavier, Spanish missionary (1506-52)

"We work for the children because children are the world's hope."

—José Martí, Cuban patriot and writer (1853-95)

"In a disparate world, children are a unifying force, capable of bringing us all together in support of a common ethic."

—Graca Machel, former Mozambique educator and champion of children's rights (b. 1945)

"Learn! No one is born full of knowledge; nor is the knowledgeable one like one who is ignorant."

—Umar ibn al-Khattab, second caliph of Islam (c. 586-644)

"You cannot teach a man anything; you can only help him find it within himself."

—Galileo Galilei, Italian astronomer (1564-1642)

"Learning is not attained by chance, it must be sought with ardor and attended to with diligence."

—Abigail Adams, wife of U.S. President John Adams (1744-1818)

"There is only one good thing in the world of which everyone thinks he has enough, and that is intelligence."

—Rene Descartes, French mathematician and philosopher (1596-1650)



"I am of those who must be taught and am seeking for teachers."

—Peter the Great, Russian ruler (1672-1725)

"Education, then, beyond all other devices of human origin, is the great equalizer of the conditions of men—the balance-wheel of the social machinery."

—Horace Mann, U.S. educator (1796-1859)

"Capacity without education is deplorable, and education without capacity is thrown away."

—Saadi, Persian poet (c. 1213-92)

"If you think education is expensive, try ignorance."

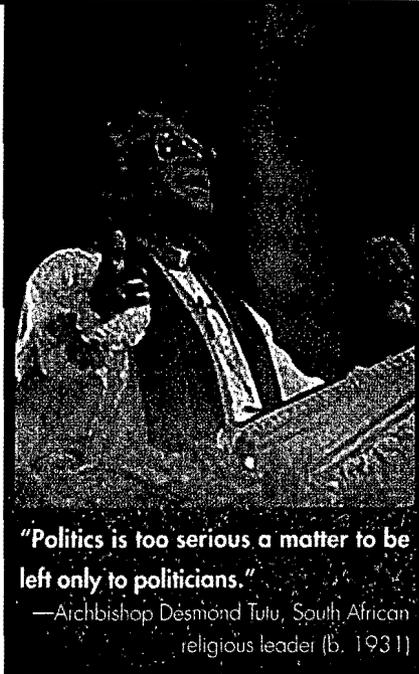
—Derek Bok, former president of Harvard University (b. 1930)

"An investment in knowledge always pays the best interest."

—Benjamin Franklin, U.S. statesman (1706-90)

"Without education, you're not going anywhere in this world."

—Malcolm X, U.S. social reformer and religious leader (1925-65)



"Politics is too serious a matter to be left only to politicians."

—Archbishop Desmond Tutu, South African religious leader (b. 1931)

"Education is too important to be left solely to the educators."

—Francis Keppel, U.S. commissioner of education (1916-90)

"People seldom see the halting and painful steps by which the most insignificant success is achieved."

—Anne Sullivan, U.S. teacher (1866-1936)

"The task of the modern educator is not to cut down jungles, but to irrigate deserts."

—C. S. Lewis, English writer (1898-1963)

"If you have knowledge, let others light their candles at it."

—Margaret Fuller, U.S. social reformer (1810-50)

"Children need love, especially when they don't deserve it."

—Harold S. Hulbert, U.S. physician (1887-1949)

"Only the educated are free."

—Epictetus, Greek philosopher (c. 60-c. 117)

"In large states public education will always be mediocre, for the same reason that in large kitchens the cooking is usually bad."

—Friedrich Nietzsche, German philosopher (1844-1900)

"What is the first part of politics? Education. The second? Education. And the third? Education."

—Jules Michelet, French historian (1798-1874)

"It is nothing short of a miracle that modern schools of instruction have not yet entirely strangled the holy curiosity of learning."

—Albert Einstein, U.S. physicist (1879-1955)

"Today the Internet is being presented as the new frontier, the new world for exploration. For those venturing into this virtual world, the charts are as incomplete as those of previous generations but, again, the thought of possible treasure lures newcomers. This world of cyberspace is a land where sirens call and dragons abound."

—Phil Buchanan, Australian educator (b. 1955)

"Education is a ladder."

—Manuelito, Navajo chief (c. 1818-94)

"What sculpture is to a block of marble, education is to the soul."

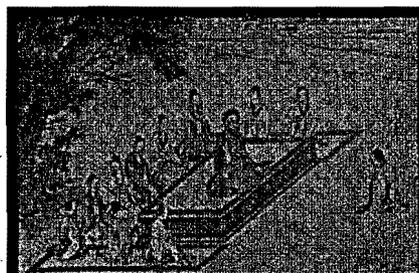
—Joseph Addison, English essayist (1672-1719)

"The supreme purpose of education, to my mind, is to inculcate the right attitudes in the hearts and minds of youth and to teach them to place first things first."

—K. G. Saiyidain (Indian author, 1904-71)

"Educate your sons and daughters, send them to school, and show them that beside the cartridge box and the jury box, you also have the knowledge box."

—Frederick Douglass, U.S. abolitionist and writer (1817-95)



"Although a man be able to recite 300 odes, if he know not how to act, of what use is his learning?"

—Confucius, Chinese philosopher (551-479 B.C.)



"Our youths love luxury. They have bad manners, contempt for authority; they show disrespect for their elders, and love to chatter in place of exercise. Children are now tyrants, not the servants of their household. They no longer rise when their elders enter the room. They contradict their parents, chatter before company, gobble up their food, and tyrannize their teachers."

—Socrates, Greek philosopher and teacher (470-399 B.C.)

"Education is the ability to listen to almost anything without losing your temper or your self-confidence."

—Robert Frost, U.S. poet (1874-1963)

"An educated man is one who has the right loves and hatreds."

—Lin Yutang, Chinese author (1895-1976)

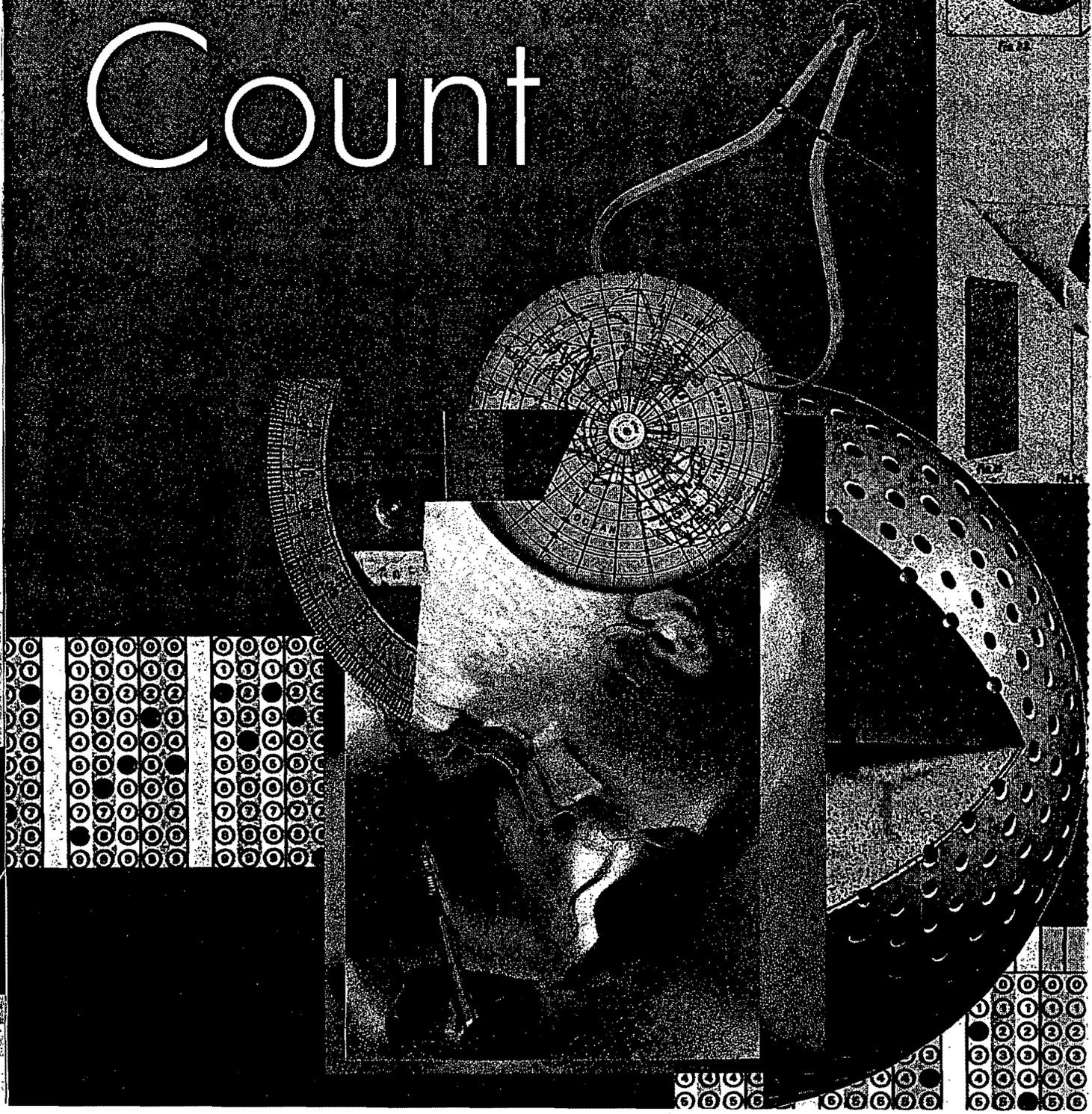
"Education is not to reform students or amuse them or to make them expert technicians. It is to unsettle their minds, widen their horizons, inflame their intellects, teach them to think straight, if possible."

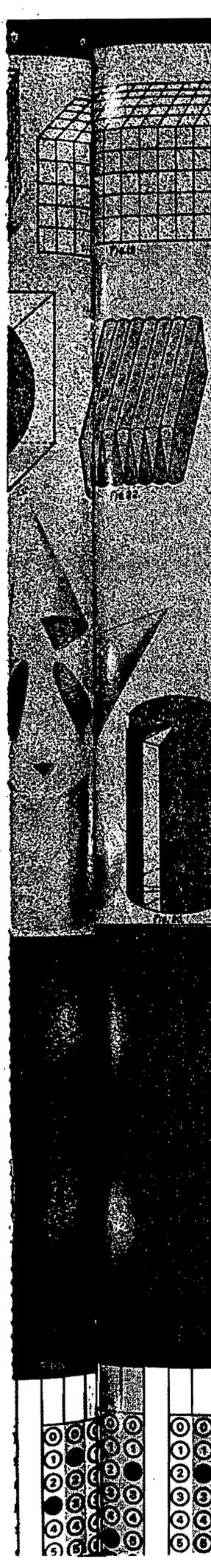
—Robert M. Hutchins, U.S. educator (1899-1977)

"Education is not merely a means for earning a living or an instrument for the acquisition of wealth. It is an initiation into life of spirit, a training of the human soul in the pursuit of truth and the practice of virtue."

—Vijaya Lakshmi Pandit, Indian political leader (1900-1990)

Tests That Count





Do high-stakes assessments really improve learning?

BY NINA HURWITZ AND SOL HURWITZ

They are tests that count, high-stakes tests that can deny promotion or graduation to students with failing scores. Schools with too many low-performing students can be exposed to the glare of publicity, placed on probation, or closed. A widening coalition of governors, business leaders, parents, and teachers—appalled that youngsters can advance through school, receive a diploma, and seek further education or a job without mastering basic skills—is promoting the use of these tests as a means of boosting standards and improving accountability in public education.

The movement is gaining national momentum. Forty-nine states have adopted performance standards for elementary and secondary education; 26 have high school exit exams in place or in process; 19 publicly identify failing schools. President Clinton is in the vanguard, calling for higher standards and a crackdown on social promotion. Last fall he urged the nation's governors, "Look dead in the eye some child who has been held back [and say], 'We'll be hurting you worse if we tell you you're learning something when you're not.'"

High standards, high stakes

High-stakes testing is forcing the debate over a fundamental question in American education—whether it is possible to achieve both excellence and equity. On one side are those who claim that tests with consequences are the only sure route to higher standards and stricter accountability. On the other are those who contend that high-stakes tests are a command-and-control instrument for "standardizing" education and punishing disadvantaged and minority children. But a more pragmatic middle position is evolving based on the experience of front-line practitioners: High-stakes testing can work with clear but limited goals, flexibility in meeting those goals, and the will to address head-

on the problems of students at risk of failure.

Texas, Chicago, and New York City and state, discussed below, are being carefully watched by educators and decision makers nationwide for both positive and negative lessons. The states are driving the high-stakes movement: Kentucky, Maryland, Massachusetts, North Carolina, and Virginia are running noteworthy programs as well.

Even as states and school districts attempt to raise standards and impose high-stakes tests, they are confronted with excessive numbers of their urban, minority, and disadvantaged students who are failing these tests. In urban districts, large-scale failure is inevitable, says *Education Week's* Ron Wolk, given the shoddy education these students are receiving. "For tens of thousands of urban youngsters, it's a kind of double jeopardy," Wolk declares. "The system failed to educate them adequately, and now it punishes them for not being educated." Schools and school districts might face punishment as well: Low scores could result in the reorganization of schools or a shift of resources to charter schools or private-school vouchers.

Parent advocacy and civil rights groups are challenging the tests on racial and equity grounds. The penalties, they claim, fall disproportionately on minority and at-risk students, who have been short-changed in their education. Meanwhile, teachers and researchers are beginning to question the tests' educational validity: Do they, in fact, improve learning?

Educators are unanimous that high-stakes tests should be aligned with curriculum and instruction—they should measure what students have been taught and are expected to know—and that teachers should be involved in the process. But only gradually are states and school districts committing sufficient resources and time to achieve proper alignment with full teacher participation.

The time lag, educators argue, makes it risky to impose consequences prematurely.

Disagreement between states and urban school districts over which test to use means students in the same grade might have to take two tests—in the same subject. Learning suffers, educators say, when teachers spend time preparing students for too many tests. “The first thing to go in a school or district where these tests matter,” says education expert Alfie Kohn, “is a more vibrant, integrated, active, and effective kind of instruction.” [See “Raising the Scores, Ruining the Schools,” *ASBJ*, October 1999.] A fifth-grade teacher in Virginia concurs: “Sometimes, when I wish I could stay longer on a subject, I have to move on to prepare my kids for the tests.”

A 1999 study titled *High Stakes* by the prestigious National Research Council sharply criticized the practice of relying solely on tests to determine promotion or graduation. Such decisions, the council argued, “should be buttressed by other relevant information about the student’s knowledge and skills, such as grades, teacher recommendations, and extenuating circumstances.” Many educators question the value of holding kids back *period*—but certainly not without a highly structured, and often costly, intervention and remediation strategy.

The growing public demand for standards with accountability has made high-stakes testing a tempting political issue. The public is fed up with low standards and courses that lack content—they want American students to be able to compete favorably with kids in other countries. Test scores provide an aura of businesslike accountability for superintendents, principals, and teachers and a stimulus for students. Initially, at least, testing seems easy and inexpensive compared with more deep-seated reforms such as hiring and training competent teachers, reducing class size, or repairing crumbling school buildings. But achieving accountability is neither simple nor cheap.

The states and school districts that have had the most success with high-stakes testing share several common characteristics. They have maintained bipartisan political support and the backing of a broad coalition of interest groups, including the business community, over a sustained period. High-stakes tests have not only raised standards but have stimulated systemwide reform. Most important, there has been a heavy investment in addressing the academic performance of the weakest students.

Turnaround in Texas

Texas is a dramatic case in point. Once considered one of the nation’s educational backwaters, the Texas public school system, according to the *New York Times*, is now viewed by educators as “a model of equity, progress and accountability.” The state’s education reforms have spanned the administrations of former Democratic governor Ann W. Richards and current Republican governor and presidential hopeful George W. Bush. In a system of 3.7 million students that is half African American and Hispanic, the scores of African-American and Hispanic students on national assessments in reading and mathematics in

1996 and 1998 outranked those of most other states, and scores on state assessments for all students have risen for the fourth straight year.

A unique feature of the Texas system is the Texas Assessment of Academic Skills (TAAS), the state’s high-stakes exam program, introduced in 1990. The tests combine clearly state educational standards with a detailed reporting of results by ethnicity and class. Scores are sorted according to white, African-American, Hispanic, and economically disadvantaged groups. Along with attendance and dropout rates, TAAS scores are used to identify a school as failing if any one of its four demographic cohorts falls below standard. “Disaggregation of scores has focused the schools’ attention on kids that were once ignored,” according to University of Texas professor U. Treisman, who is director of the Charles A. Dana Center in Austin. Texas is gradually raising the passing bar to 50 percent for each cohort from the original 20 percent.

Until recently, the high stakes associated with the TAAS have consisted almost entirely of public disclosure of school-by-school test results, a process Gov. Bush calls “shining a spotlight of shame on failure.” The ratings, published on the web to identify schools as exemplary, recognized, acceptable, or low performing and provide strong incentives to improve for adults and students alike. For example, superintendents, principals, and teachers find it hard to get jobs if they come from failing schools. Although low-performing schools are bolstered by additional financial support, they are rarely closed. “There are a great ideas on what to do with really problematical schools,” says Treisman.

Elementary and middle-school students are tested in grades three through eight on various combinations of reading, writing, and mathematics, with science and social studies added in the eighth grade. In response to political pressure, Texas will move to end social promotion by 2003. Hoping to avoid widespread retention, the state has instituted the Student Success Initiative, an early-intervention program, starting with the current year’s kindergarten class. A skeptic on retention, Treisman cautions that research on the dropout problem indicates that “being overage in your class has the single highest correlation for dropping out and is twice as high as for any other factor, including race.” Also, as pressure mounts to pass the TAAS, state officials have become increasingly concerned over outbreaks of alleged cheating.

There is wide agreement that Texas high schools have not shown as much improvement as elementary schools. However, the state plans to beef up the content of the 10th-grade exam and move it to the 11th grade and to allow substitution of end-of-course exams in core subjects. The present 10th-grade exam is the subject of a lawsuit by the Mexican-American Legal Defense Fund in the U.S. District Court in San Antonio, which claims the test discriminates against minority students. Gov. Bush counters this claim: “Some say it is a racist test,” he told the *New York Times* last summer. “I strongly say it is racist not to test because by not testing we don’t know, and by not knowing we are just moving children through the system.” The or

come is sure to have an impact on other states, researchers agree.

Remarkably, the state's educational resurgence has occurred while expenditures remained below the national average: In 1998-99, Texas spent \$5,488 per student compared with the national average of \$6,407. Striking, too, is the autonomy that Texas gives its principals and teachers as long as test results remain positive. Bilingual education, for example, is a local option. However, scores for Spanish-speaking and special education students must now be included in overall ratings to ensure more accurate results.

Success in Chicago

Just as Texas has drawn raves for educational attainment at the state level, Chicago, with an enrollment of 431,000 students, has become the promised land for city reformers. Mayors, superintendents, and educators have flocked there to study the remarkable turnaround orchestrated by chief executive officer Paul Vallas, formerly budget director under Mayor Richard M. Daley. With no previous experience in education, Vallas has performed what many consider an educational miracle in a school district that U.S. Secretary of Education William J. Bennett in 1987 called "the worst in the country." Vallas achieved credibility largely through the selective but determined application of high-stakes testing. Scores on Chicago's performance benchmark, the Iowa Test of Basic Skills in Reading and Mathematics, have risen for the fourth straight year.

When Vallas took charge in 1995, an earlier reform effort, which stressed decision making by local school councils, had virtually hit bottom. With high truancy, low standards, and rampant grade inflation, Vallas declared, "there was wide agreement that the earlier reform initiative had failed." Vallas exploited public dissatisfaction with the previous reform, while drawing grassroots support from a network of parents, community groups, foundations, and universities to fashion a new strategy.

Fundamental to his success was the solid backing of Mayor Daley and Gery Chico, president of the Chicago School Reform Board (successor to the former elected school board), whose members were all mayoral appointees. Vallas forged a close working relationship with Tom Reece, who heads both the Chicago Teachers Union and the Illinois Federation of Teachers, and together they have succeeded in avoiding strikes and confrontations by building communication and trust. In addition, he won points with the public for his skills as a financial manager by stamping out waste, ending deficits, and securing state funds for building and renovating schools.

Three years ago, Chicago gained national attention as the first big-city school system to end social promotion. Students who don't meet minimum standards on the Iowa Tests are at risk of retention, but the passing bar was set low at first to avoid massive failure and is only gradually being raised.

Chicago's promotion gates kick in for students in grades

The growing public demand for standards with accountability has made high-stakes testing a tempting political issue.

three, six, and eight. For those who fail, the city's mandatory Summer Bridge Program, staffed with experienced teachers, provides a scripted curriculum from the central administration with hour-by-hour guidelines. University of Chicago professor Melissa Roderick, a member of the Consortium on Chicago School Research, believes the program goads parents, students, and teachers to work harder to avoid retention. "Students love Summer Bridge," says Roderick, because they know it helps them. Of the estimated 25,000 students who attended summer school last year, two-thirds moved to the next grade. Chicagoans call the policy "retention plus" because it comes front-loaded with ample resources for intervention and remediation. "Retention is a last resort," Vallas maintains.

Chicago also provides tutoring during school, and in an after-school Lighthouse Program (with supper included), for students who fail. Cozette Buckney, the system's chief education officer, shares the prevailing view of education experts that students should not merely repeat the same curriculum once they are held back. "You must teach them differently, use different materials—give them a different experience," she says.

Finally, if students have not passed the eighth-grade test by age 15, they move to "academic prep centers" that offer small transitional classes and intensive test preparation, where expenditures per pupil are one-and-a-half times those for high schools. Most students move on to high school after one year, although some teachers believe the centers accentuate problems of self-esteem and increase the tendency to drop out.

So dazzling is Chicago's success in the lower grades that outside observers have hardly noticed that real achievement stops at the high school door. "In the high schools, we have been at a loss," Buckney admits. Standards remain low, and there is widespread disengagement of students, a weak curriculum, and meager support services. Half of the city's ninth-graders fail two or more courses.

The Chicago Academic Standards Exams (CASE), which are end-of-semester high school tests in core subjects, are currently being upgraded. Teachers now receive detailed content guides from the administration but complain about the rigidity that the guides impose on their teaching. Although most teachers allow the exams to count for only 15 percent of the semester grade.

a biology teacher contends that “the tests shape what I teach, what order I teach in, and how long I spend on each subject.”

High school teachers have been more resistant to control from the central office than elementary school teachers, according to Vallas. “They view themselves as college professors—they’re more set in their ways,” he complains. Last year George Schmidt, an activist teacher, published parts of the CASE tests in protest, and students at top-rated Whitney Young High School boycotted the tests. Vallas dismisses such opposition, saying, “There is enough to be irritating but not enough to delay reform.”

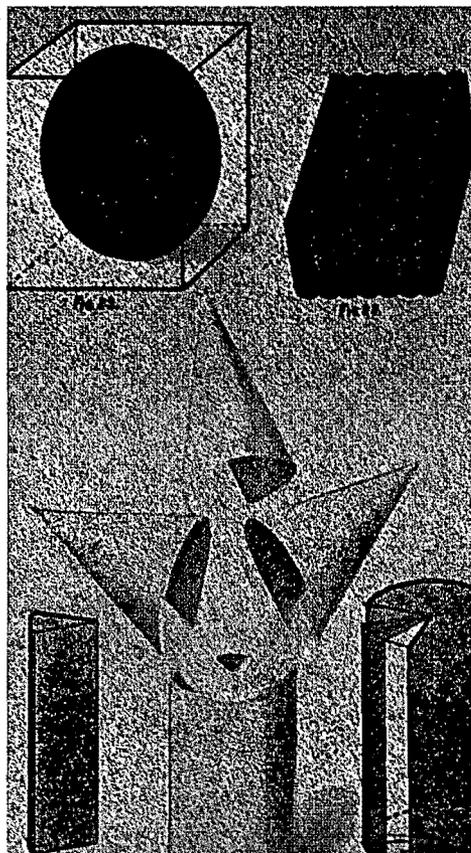
Recent efforts to close down and reconstitute Chicago’s worst high schools have proved unsuccessful, and discharging low-performing administrators and teachers has been difficult. “The burden of proof [on the school administration] in removing failing teachers is pretty strong,” Vallas admits.

Vallas is seeking to garner support for Chicago’s high-stakes tests by allowing waivers and retesting. He is also identifying at-risk students early and conducting special programs for pregnant teens and teen mothers. “Good attendance, behavior, and grades” can help students get promoted, he says. The time will come, he predicts, when tests will become a diminishing factor in promotion decisions.

Disappointment in New York

Chicago’s success in strengthening standards and ending social promotion in the early grades contrasts sharply with New York City’s recent dismal experience with high-stakes testing. New York’s gigantic scale—it is the nation’s largest school system with 1.1 million students—and the fractured relationship between the schools chancellor and the mayor have vastly complicated attempts to impose high stakes on state and city tests. Unlike Chicago, where Paul Vallas and Mayor Daley work in blissful harmony, New York City’s Schools Chancellor Rudy Crew has recently been at loggerheads with Mayor Rudolph W. Giuliani.

The past year has brought high-stakes testing to the boiling point. After months of cramming and intense pressure on students, teachers, and parents, the news came in May that 67 percent of New York City students had failed the state’s new and more demanding fourth-grade language arts test. The test was given over three days and included passages to be read for



High-stakes tests are transforming the education landscape, and lawmakers and educators are learning to navigate in uncharted terrain.

comprehension and answered in essay form. Stunned by the low test scores, the mayor proposed the removal of principals from the bottom third of all city elementary schools and called for a major management shake-up. State Education Commissioner Richard P. Mills recommended summer school for all students who failed the test.

Then came the disappointing results on year-end city tests, administered and graded by CTB/McGraw-Hill, which showed only 44.6 percent of students reading at grade level, a five-point decline from the previous year. Mathematics scores were even lower, falling 10 percent. After this second dose of bad news, the mayor prescribed even stronger medicine. Impressed by Chicago’s example, he called for abolishing the semi-independent board of education and placing the schools under his own control.

After constant badgering from the mayor, Crew responded in mid-June with a hastily arranged mandatory summer school program starting in July for 37,000 third-, sixth-, and eighth-graders who had scored at or below the 15th percentile on the city’s standardized reading test and the 10th percentile on the mathematics test, both taken in the spring. Students failing the tests a second time would be held back.

The summer-school program was plagued with problems. For six weeks the schools were forced to cope with thousands of youngsters who needed to pass the city tests to avoid retention. Many teachers were handicapped by a lack of student records and by inadequate course materials, and school buildings were stifling from a record-setting heat wave. Although paid at a lower rate than their school-year salaries, many teachers had to buy their own materials and bring fans from home. In some instances, students who were supposed to take on the mathematics test were drilled mainly for the reading test. Chancellor Crew’s pride in announcing that 64 percent of the students passed the tests was soon dampened when scoring errors by CTB/McGraw-Hill revealed that more than 8,600 students were sent to summer school by mistake. Lack of accurate summer-school attendance figures cast further doubt on the number of students who would be retained. The Puerto Rican Legal Defense Fund and Advocates for Children, a nonprofit legal services organization, have threatened court challenges, citing late warning of the new requirement to attend summer school and the use of a single test to determine promotion.

New York City’s problems are likely to be exacerbated if

Commissioner Mills' determined belief in high-stakes testing as a means of raising standards for the state's high schools. In a program that is unique in the nation, Mills and the State Board of Regents are requiring that by 2003, all students will have to receive a passing grade of at least 65 percent on the state's tough new Regents examinations in five subjects—English, mathematics, science, global studies, and American history—before they can graduate. Currently less than a quarter of New York City students qualify for a Regents diploma.

Turning aside his critics, Mills contended in a *New York Times* interview last summer that without high standards, "You simply decide in advance that some students don't have access to the good life. They don't have access to jobs, they don't have access to enriched curriculum and everything that goes with it."

Mills understood the need for a structured program of remediation and support for those who fail the Regents tests. To this end, he proposed a \$900 million program targeted toward poor districts, but Gov. George E. Pataki's budget came nowhere near to providing that amount. Last spring state lawmakers—under pressure from parents, teachers, and school administrators who feared widespread failure on the Regents tests—argued for scaling down requirements. In October a consortium of parents and educators at 35 New York City alternative high schools asked legislators to compel the Regents to exempt its students from the new English Regents exams. With public opposition rising, it is doubtful the Regents will have enough public support to sustain such an extensive testing program with such high stakes over the long term.

As the results of last year's more rigorous six-hour, two-day English Regents tests were being released, accusations of deceptive scoring on the essay questions began to surface. According to a Harlem high-school teacher, "I never would have given points in a regular class for the kind of answers we were getting on those essays."

Many teachers had never seen the state's new standards; nor had they been trained to teach courses to the level of the Regents' demands. "No business or military organization would do that kind of campaign without adequate training," Thomas Sobol, former state education commissioner, asserted at a meeting in Purchase, N.Y., last fall. For New York City's students, the stakes are overwhelming and probably unrealistic.

Making high-stakes testing work

High-stakes tests are transforming the education landscape, and lawmakers and educators are learning to navigate in uncharted terrain. Conditions and requirements vary state-by-state, and progress in meeting the new standards requires patience. But some early lessons can be drawn from states and school districts that are beginning to achieve success:

■ **Make sure that learning—not testing—is the goal.** "Are we teaching for testing or teaching for knowledge?" a senior administrator asks. Tests can be important in identifying weaknesses. But too much testing in too many subjects overwhelms teachers, drains resources from enriched educational programs, stifles creativity, and increases cheating.

■ **Give disadvantaged students special assistance.** High-stakes tests can be a powerful tool for raising standards for at-risk students, but only if resources are reallocated to schools that serve them. And the testing program must be held accountable for ensuring that the tests are reliable, fair, and free of cultural bias.

■ **Set failure rates at a realistic level.** Most schools lack the resources and capacity to absorb masses of failing students in after-school and summer-school remediation programs and to conduct programs for students who repeatedly fail. But setting failure rates too low damages credibility in the system's standards. The right balance will vary according to circumstances, but finding it is crucial.

■ **Invest in a wide range of educational reforms—not just tests.** Tests don't work in a vacuum but in an environment that supports systemwide reform. Tests should be part of a program that encourages early childhood education; the recruitment, training, and development of capable teachers; smaller class size; and safer buildings.

■ **Make retention a last resort.** Most studies show that retention does more harm than good. Frequent failure erodes self-confidence, and students who are retained have a higher probability of dropping out. If retention helps at all, it does so only when students are supported by innovative learning strategies. Decisions to deny promotion should not be based on a single test and should involve the teacher.

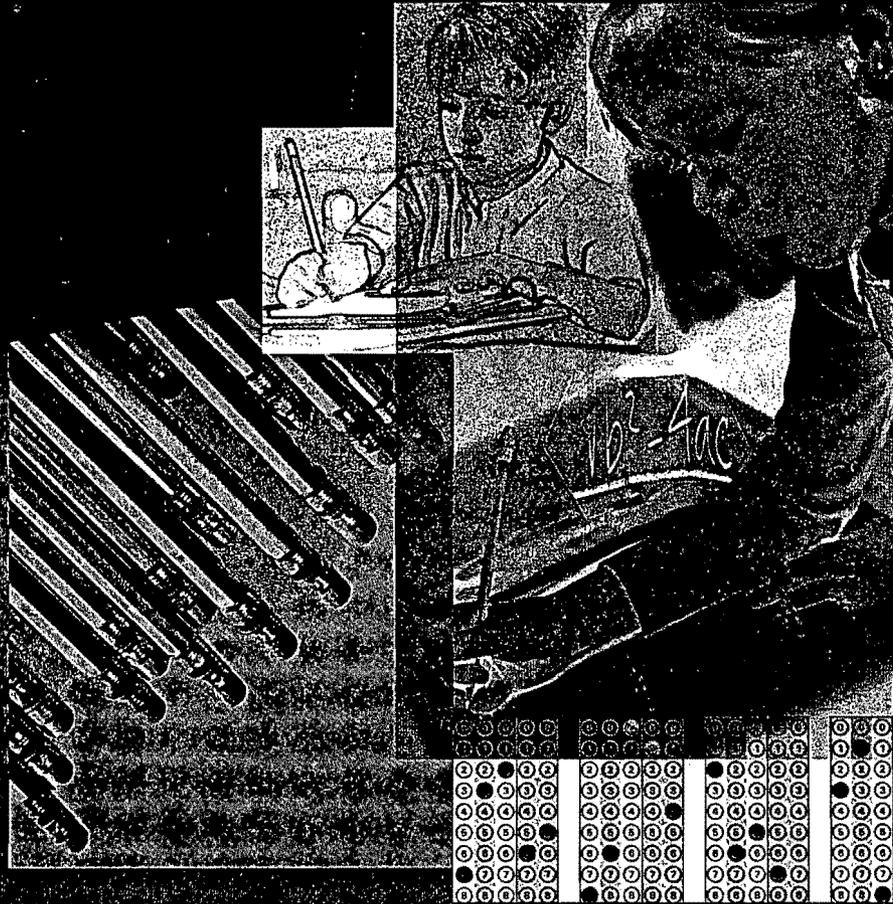
■ **Use publicity to force improvement.** School rankings draw attention to the weakest schools and can be used to drive decisions regarding school reform, reorganization, or closure. School officials have an obligation to interpret test results to the public consistently and accurately and to be forthright about problems in the system.

■ **Focus on urban high schools.** Tests can be effective in raising standards but only if problems of school climate are addressed. Expect high school exit exams to be challenged in the courts by minority groups. Excessive testing narrows curriculum choice, and the need for remediation may lead to *de facto* tracking and high dropout rates.

■ **Prepare for the long haul.** It is the rare state or school district that gets high-stakes testing right the first time. Success takes time and requires experimentation. Be ready to adapt, adjust, and compromise in order to achieve long-term success.

On balance, high-stakes tests that are well-designed and carefully administered appear to be working, at least in the lower grades. But if their benefits are oversold and their dangers ignored, disenchantment could lead to diminished support for public education. If, on the other hand, they call attention to failure and encourage strategies to ensure success, they could stimulate a long-overdue educational renewal for the nation's neediest students.

Nina Hurwitz was a high school social studies teacher in Westchester County, N.Y., for 23 years. Sol Hurwitz (solhurwitz@aol.com) is an education consultant and freelance writer in Rye, N.Y.



None of the Above

BY FREDERICK M. HESS
AND FREDERICK BRIGHAM

Standardized accountability systems are steadily gaining popularity across the United States. Ambitious programs have been launched and implemented in states as diverse as Texas, Ohio, and Virginia. These programs—from the Texas Assessment of Academic Skills (TAAS) tests to Virginia's Standards of Learning (SOLs)—are based on the notion that there is a specific body of knowledge that students must master. This belief has long characterized Western European and Japanese schooling, but it marks a radical departure from the traditional American approach to schooling.

Rather than attempt to assess whether this change is good or bad, we are interested in how it affects educational practice in schools and classrooms. It is our belief that these accountability systems cannot easily coexist with the traditional administration and culture of schooling—one or the other has to give.

Standardized accountability systems are predicated on the idea that all students will learn a predetermined body of knowledge to a particular level of accomplishment. The American version of accountability is actually more demanding than the versions in other industrialized democracies, because the U.S. debate assumes that all students will meet uniform standards. Nations such as France or Japan, on the other hand, tend to use their testing regimes to sort students into vocational tracks; the

very existence of these tracks presupposes that not all students will achieve at a high level on the tests.

A double-edged sword

The promise of high-stakes tests is that they can set a clear and challenging hurdle for students and for schools. High-stakes testing that has real consequences can motivate significant educational improvement. In doing so, however, such testing puts the state in the business of labeling significant numbers of students as "failures." High-stakes tests also put state officials in the business of making formal decisions about what academic content is more or less valuable for students to learn.

Setting high standards means that some students will fail to meet those standards. And, as the test results show, those who fail are likely to be disproportionately poor and African American. Students with disabilities also tend to fare poorly on such tests, underscoring the tension between our desire to offer all children an opportunity to excel and our desire to hold all students to high standards. Most of these students might benefit in the long run, because communities that have been ill-served by schools demand school improvement. But long-term benefits do not compensate those students who lose out in the short run.

In the short run, schools will deny diplomas in disproportional

The promise and peril of high-stakes testing

tionate numbers to poor, minority, and special-needs students. Even dramatically improving schools in disadvantaged communities is unlikely to ensure passage rates comparable to those in well-off communities, as data consistently have shown that advantaged students benefit educationally from the socioeconomic status and educational attainment of their families and friends.

Are voters and policymakers willing to sustain practices that make children from certain backgrounds less likely than others to receive a high school diploma? In truth, that is not so different from the situation today. Today, minority students and poor students are much more likely to drop out, to receive a diploma of little value, or to not attend college. Now, the barriers are generally informal and invisible, but standardized evaluation lends the stamp of civil approval to these inequities.

This kind of state-sanctioned sorting is an accepted part of life in many industrial democracies, such as France and Japan. Faring poorly on high-stakes assessment in these nations means being slotted into a less prestigious career track, but government policy has consciously sought to reduce the income and lifestyle difference between career tracks. Part of the social contract in these nations is a promise of relatively comfortable working conditions for those who have not cleared the hurdles.

Historically, however, Americans have been less comfortable with the government sorting out individuals and then actively seeking to moderate the rewards for success and penalties for failure. Here, opportunities are less dependent on government decisions; here, the rewards for success are larger, and so are the penalties for failure. Our society is filled with people who have achieved wealth and success despite unimpressive educational track records. Even in our relatively wide-open economy, however, at least some postsecondary training is generally necessary if one is to have a real shot at wealth and success. Consequently, it is particularly significant when the government denies some children diplomas, thereby limiting their future opportunities. Even though the blame for these students' failure is partially due to their schools or their social circumstances, the full penalty will be borne by the students who do not receive diplomas.

Do we have the will to make these children pay this price?

Weighing the costs and benefits

The promise of high-stakes evaluation turns on the balance among several costs and benefits, of which the following are particularly important:

■ **Increased equity.** Statewide assessments have the potential to increase the equality of educational opportunities available to students across school districts. Standardized assessment will dramatically reduce the inequities that result from different curricula being offered in different schools or even within different classes in the same school. As E. D. Hirsch has pointed out, nonuniform curriculum is a problem in a highly mobile society. Students can move from the upper quartile to the lower quartile of their school simply by moving 20 miles down the road to a school that follows a different curriculum. Also, students in different classes can receive varying amounts of instruction on different aspects of the curriculum, depending on their teachers' interest and expertise. It is unreasonable to expect all teachers to devote similar amounts of instruction to the different aspects of the curriculum without some sort of external guide. Statewide assessment standards can provide one source of external guidance for teachers and students.

■ **Clarity and focus.** Because statewide assessments can bring clarity and focus to the curriculum, teachers and students can have a more solid understanding of what is successful teaching and who are the best students. Under statewide assessment systems, good teachers are those whose students attain high scores on the tests, and good students are those who attain the high scores. Classroom grades, attendance, and nonacademic

aspects of school performance such as citizenship do not count—only proficiency on the standards assessment.

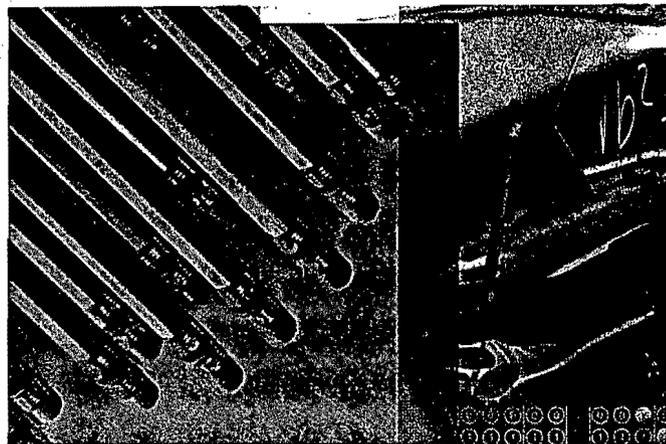
In Virginia, which announces the format of its tests (mostly multiple choice), a savvy teacher can identify the instructional objectives most likely to appear on the test and focus on those items. There are few tangible benefits for teachers who focus on items outside the tested curriculum or on items within the curriculum that are unlikely to appear on the test. In short, when administrators and the public use the tests as a barometer of educational performance, teachers will focus their efforts on the areas most likely to be measured.

■ **Efficient use of resources.** A more focused curriculum leads to more efficient use of resources. Because most teachers will be teaching the same thing to most students, many costs might be reduced. Purchasing agents will be in a position to select materials that are explicitly related to specific instructional objectives or parts of the statewide assessments. Professional development can be much more narrowly conceived and directed.

■ **Evaluations of school personnel.** Because statewide assessments focus on specific curriculum objectives, teacher evaluations will focus on how well students test. Teachers are either working on the statewide assessment objectives or they are not. Principals and supervisors can easily check plan books and students' work for evidence of adherence to the implicit curriculum outlined by the statewide assessment. Student scores on the tests provide an easy, concise, and nearly irresistible index of teacher performance. In some schools, teachers have been told, "Your students' test scores are being monitored. If they do not improve, you will be fired." It is hard to argue that teachers are unaware of the standard by which they are being judged. The message is clear: Improve the test scores of your students above all else. This is obviously an effective motivating device in some ways, but it might not produce the kind of classrooms or schools that we really desire.

When great emphasis is placed on any particular educational outcome, the curriculum and the actions of school personnel become aligned with that goal. Either the district's scores on the statewide assessment are considered adequate or at least improving or they are not. For students, eligibility for a high school diploma is often tied to performance on statewide assessments. For schools, funding, accreditation decisions, and contracts for personnel are often tied to students' performance on statewide assessments. A similar focus can be placed on principals and central office administrators. Indeed, some school districts have adopted a policy that calls for dismissing the administrators in schools whose students turn in inadequate performance. As a result, teachers in Ohio and Texas have informed us of schools where a substantial part of the curriculum now revolves around test preparation. That means that the items on the test and the way they are constructed are critically important, because they will to a great extent determine what teachers and students do in school.

■ **Limited local decision making.** When state funding allocations are tied to student performance on statewide assess-



ments, a *de facto* state curriculum is created. Department of education officials and test creators can impose their views of curriculum content, emphasis, and timing on local districts with relatively little consultation with their constituents. Ironically, the statewide testing movement is taking place at the same time as states are experimenting with charter schools, voucher systems, and increased acceptance of home schooling. These experiments encourage increased diversity and choice, while the statewide assessment movement encourages greater homogeneity among educational programs. The choice offered parents is to follow the implicit curriculum of the statewide assessment or strike out on one's own. Planning and innovation at the community level appear to be omitted in both of these movements.

■ **Narrow curriculum.** Statewide assessments are not meant to suggest that only what is on the test is important, but many schools have interpreted them this way. Teachers and administrators have reported that programs for the arts, vocational education, and physical education are sharply limited by these decisions. A northern Virginia superintendent told us she was considering a recommendation to move the arts to a weekend program, eliminate vocational education, and turn physical education into an honor system in which students would sign pledges that they had engaged in calisthenics to prepare themselves for the President's Physical Fitness Test.

Another school official in central Virginia said her district had adopted a policy that only activities and projects focused on the state Standards of Learning test would be allowed in its schools. The SOL assessments cover four areas: language arts, mathematics, social studies, and science, with accompanying standards for technology across these four areas. Although these areas are critically important, they are not exclusively important. Emphasizing these areas to the exclusion of other areas of curriculum is not likely to produce the kind of well-rounded citizen that life in the 21st century demands.

An example of how a test shapes curriculum with unfortunate results can be found in the medical field. In a 1997 paper, Judith Koenig and Andrew Wiley described a pattern of course taking called the "premedical syndrome." Because the Medical College Admission Test (MCAT) focused primarily on four areas of science, premedical students focused their coursework and extracurricular activities narrowly on the sciences to the exclusion of studies in humanities and social sciences. Contemporary medical practice, however, requires a broader background to support communication, problem solving, and data analysis. Researchers also discovered that simply testing well

on the MCAT was an insufficient predictor of successful medical school performance and subsequent practice. Consequently, the test and its role in medical school admissions were carefully examined and revised. We doubt that the content and cut-off criteria for statewide assessments have been subjected to that kind of rigorous evaluation. The potential for state tests to shape the implicit curriculum is clearly as strong as if not stronger than that of the MCAT because statewide assessments influence curricular decisions across entire school districts rather than the decisions of a few individual students.

■ **Favored content.** High-stakes testing elevates some content above other content and focuses schools on teaching the kind of material the tests measure. The tests require us to decide what knowledge and skills we value most. The benefits we glean by creating clear content and performance expectations are obtained at the expense of other material.

Large-scale testing generally emphasizes content knowledge rather than higher thinking or developmental skills. If there are strong disagreements about what students should know, then the need to decide what to include in the tests increases the significance of those arguments. This can produce fierce conflict, particularly regarding what skills we want schools to develop and what knowledge we consider most important in politicized subjects like history, social studies, and the humanities. Classroom teachers might find themselves actually insulated from responsibility for what they teach. Some teachers might find this confining, while others might be relieved to be removed from the hot seat.

Over time, if the assessment is institutionalized, conflict about what should be on the test is likely to become muted. People will become much more concerned with how students fare and with preparing them to take the test—a prospect that offers scant comfort to those concerned about what is missing from the test.

A special case

Special education offers a particularly apt illustration of the tensions between high-stakes testing and other assumptions underlying American public education. At the same time that schools are required to raise standards by implementing rigorous testing procedures, they are also required to include many students who have historically been excused from participation in assessments.

The 1997 revision of the Individuals with Disabilities Education Act explicitly calls for students served in special education programs to participate in districtwide and statewide assessments except in specifically defined circumstances. This requires schools to take greater responsibility for the achievement of students with disabilities, but it also encourages administrators and teachers to look for loopholes to excuse more of these students from high-stakes assessments to keep average scores from declining. For instance, many Texas schools are said to use the English-as-a-second-language provisions to exempt large numbers of marginal students from the TAAS assessments.

In addition, the current trend toward inclusion has resulted in increased numbers of students with disabilities being served in general rather than in special education classes. Teachers are required to raise the test scores of the students in their classes at the same time they are required to accept more students with historically low achievement and test scores. Many teachers see themselves in a bind, unable to accommodate this increased diversity and improve the test scores of their students at the pace demanded by these two competing movements. Asking school personnel to negotiate two sets of conflicting demands is likely to result in frustration and disappointment.

Minimizing the costs

Rigorous adherence to specified content and standards is likely to improve the state educational system as a whole but inflict costs on some readily identifiable and vulnerable populations. Long-term educational improvement is likely to emerge as schools begin to focus on a clear set of outcomes and pay less attention to various distractions. Additionally, clear measures of performance might prove valuable in targeting resources, focusing public attention on school quality, assessing teacher performance, and fostering respect for the professionalism of demonstrably effective teachers and administrators. But the potential costs of high-stakes statewide assessment are clear.

School leaders must fight to ensure that the assessments developed for their state promote the benefits and minimize the costs. Frequent changes in content or in standards will discredit the standards, creating cynicism and apathy, and will encourage educators and policymakers to focus on the gamesmanship of standards. In the absence of focused leadership, standards are likely to waste resources, hurt teacher morale, distract attention from the hard work of improving schools—and take a disproportionate toll on poor children, minorities, and students with disabilities.

Frederick M. Hess (fmh3x@virginia.edu) is an assistant professor of education and government at the University of Virginia, Charlottesville. Frederick Brigham (brigham@virginia.edu) is an assistant professor of education at the University of Virginia.

References

- Farmer, Robin. "Disparity in Tests 'Alarming.'" *Richmond Times-Dispatch*, May 31, 1999. p. A1.
- Hirsch, E.D. Jr. *The Schools We Need and Why We Don't Have Them*. New York: Doubleday, 1996.
- Koenig, Judith A., and Andrew Wiley. "Medical School Admissions Testing," in Ronna F. Dillon, ed. *Handbook on Testing*. Westport, Conn.: Greenwood Publishing, 1997.
- Reeve, Douglas B. *Making Standards Work: How To Implement Standards-Based Assessments in The Classroom, School, and District*. Denver, Colo.: Center for Performance Assessment, 1998.