

# Test Mess 2: Are We Doing Better a Year Later?

The rage for (and increasingly against) testing and accountability shows no signs of abating. Mr. Goldberg updates his own assessment of developments on the testing front.

BY MARK GOLDBERG

**I**N JANUARY 2004 I published an article in the *Kappan* titled “The Test Mess.” In it, I examined how we were doing with state and federal accountability and tests. It was clear at that time that tests and accountability were not going to disappear — or even diminish — as the central mechanisms of the national effort to improve education. It was also clear then that many of the efforts being made were facing some doubt or encountering some difficulty. This year, it is clear that most of the issues in play have become even more complex.

First, much more media attention has been paid to testing — and in particular to the No Child Left Behind (NCLB) Act and its requirement for Adequate Yearly Progress (AYP) on state tests. Second, the number of objections to various aspects of testing has increased dramatically, and these objections call into question everything from the tests themselves to the lack of funding to support federal mandates. Third, many cities and states, and even parents in some cases, have found interesting and creative ways to circumvent the tests or their results. Finally, the U.S. Department of Education (ED) has compromised on some

of NCLB’s requirements and has faced considerable pressure to back down even more.

In this article, I take up roughly the same five questions I raised last year and examine the interesting developments that have taken place since. I say “roughly” because I have refocused some of the questions slightly in order to capture more accurately the events of the past year.

**1. Are the tests and the testing movement as currently conceived worthwhile in the first place?** Although some states have made an effort to improve their tests in the last year or two, the vast majority of test questions remain narrow and often trivial, largely for ease of scoring, which holds down the cost. Many of the test items are of questionable value, asking for information that successful adults forgot long ago. As Dave Posner points out, “The kinds of problems that can appear on a standardized test are, of course, quite limited in form and complexity, as the student is allocated only a minute or two to complete each one.”<sup>1</sup>

Teachers frequently complain that the amount of time devoted to test preparation for narrow-gauge tests eliminates time to train students in how to approach a problem or think through an issue. Creativity, perseverance, ability to work in a cooperative group, initi-

---

MARK GOLDBERG is an education writer, book editor, and consultant. He is currently completing his fifth book, *A Handbook for Administrators*. He can be reached at [Mark12738@aol.com](mailto:Mark12738@aol.com).

ative, integrity, discipline, performance excellence, unusual focus on a significant problem, flexibility, and other laudable characteristics are rarely taught and scarcely counted toward the judgment of students' performance, yet these and similar characteristics are the primary criteria in the annual reviews of some of our finest managers in politics, business, and all forms of culture and education.

Since nearly all the states use some form of standardized tests, are they doing all they can in the way of reporting useful data? W. James Popham talks about the "instructionally insensitive" tests that most states use.<sup>2</sup> These are tests that do not help anyone figure out how to bring about improvement. If the tests told you precisely what areas of, say, reading needed to be improved, that could be of some help. If they told you just where Jennifer or Juan needed more instruction, that would be even better. But as one third-grade teacher in Maryland complained, the only information she got back was the percentages in each group of students who failed reading. There was nothing about whether the problem was vocabulary or comprehension, "nothing at all about the particular skills areas in which an individual student had weaknesses."<sup>3</sup> Because the data weren't broken down, the information was essentially useless to a classroom teacher.

Susan Neuman served in ED as assistant secretary for elementary and secondary education until January 2003. She is now back at the University of Michigan as a professor. Neuman supported NCLB while she was in the Administration and continues to support much of it today, but her support is not nearly as strong or as broad, now that she's seen the law's unintended consequences. She understands that many schools simply do not have the resources required to bring about serious improvement and so "we are creating a fantasy" that all children can achieve. Neuman goes on to say, "We all know children are being tested too much. Let's be honest."<sup>4</sup> She concedes that too many schools are failing to make AYP because of a statistical anomaly and not because of any real measure of progress. For instance, a school that moves its English-language learners from, say, the 30th percentile to the 40th percentile can still be labeled as having failed to make AYP if the passing score is set at the 60th or 70th percentile.

The essential problem with narrow-gauge standard-

ized tests will always be their inherent limitations. As the directors of the National Center for Research on Evaluation, Standards, and Student Testing say, unless and until the tests "seriously incorporate local and/or classroom assessment (performance assessment), accountability systems cannot assure adequate reliability and validity for individual decision-making purposes and cannot provide necessary information to support

*Most of the national testing companies do not recommend that their tests be the sole gauge for high-stakes decisions.*

student learning."<sup>5</sup> A year's or a semester's education is just too complex to reduce to a single test, and one test can't tell us much about a student. In fact, most of the national testing companies do not recommend that their tests be the sole gauge for high-stakes decisions. In New York City, for example, where a single test will determine which students are retained in the third grade, "even Harcourt Assessment, Inc. and CTB McGraw-Hill, the companies that make the tests, have cautioned against using test scores as the sole factor in such decisions."<sup>6</sup>

**2. Are we really going to label thousands of public schools — some long considered excellent — as failing, in spite of growing protests?** Approximately 26,000 of the nation's 91,400 public schools failed to make AYP in the 2002-03 school year. The Connecticut Education Association estimates that, if no changes are made in NCLB, "More than 90% of Connecticut elementary and middle schools won't meet federal education standards in 10 years."<sup>7</sup> Of course, the states will not tolerate that outcome.

Beginning in December 2003, under considerable pressure from policy makers, parents, and educators around the country, ED was forced to tweak NCLB's requirements by using more generous interpretations of how to treat the scores of disabled students. In February 2004, the rules on English-language learners were loosened, and in March 2004, states were allowed to average test-participation scores over a two- or three-year period. Beyond that, there is a negotiation game that can be played. For instance, in Lake Alfred, Florida, principal Eileen Castle complained that her special education teachers did a commendable job, but many

of their 34 students still failed the test. In Florida, a special education subgroup is defined as having 30 or more students, which meant that Castle's school failed to make AYP. Now if Castle "could pick up Lake Alfred and move it to Texas, her failed special education subgroup would disappear. Texas officials negotiated a larger minimum subgroup size of 50 with federal officials."<sup>8</sup> By the way, Texas also negotiated for a lower math standard (33%), while those beleaguered folks in Florida still have to meet a 38% standard in math. How long do you suppose it will take state officials elsewhere to catch on to the negotiating game?

Not very long, it seems. According to Lynn Olson, "more than 40 states have notified the U.S. Department of Education that they want to make changes in their accountability plans approved last year," particularly since they do not believe that some of the schools labeled inadequate actually are sub par.<sup>9</sup> Many states have been working hard on upgrading their curriculum and accountability systems for a decade or more. "In North Carolina, more than 90% of schools are meeting the state growth goals, but only 47% of schools made AYP. Of those that failed, 283 missed the federal goal because one subgroup fell below its proficiency level."<sup>10</sup> From Arizona to Michigan to Alaska and from Kentucky to Maryland to California, the federal government is getting complaints about the unfairness of NCLB, complaints ranging from interference with states' rights to inadequate funding, too many subgroup requirements, inappropriate and unwarranted rigidity over requirements, and overreliance on a small set of standardized tests.

Maryland, for one, is searching for methods to make it easier for schools to meet NCLB standards, mostly by making subgroup changes and interpreting the tests somewhat differently. In Michigan, officials simply inflate state self-evaluation scores to keep schools from failing. In Arizona, a bipartisan group of lawmakers has recommended that the state simply opt out of the law's accountability requirements. Oregon has decided to allow Spanish and Russian speakers who have not mastered English to see test questions in their native language. The NCLB law is more than 1,000 pages in length and contains many complex and somewhat ambiguous provisions and arcane formulas. Former Secretary of Education Rod Paige consistently blamed opposition to NCLB on liberal unions, but Scott Young, an analyst with the National Conference of State Legislatures, which represents 7,500 legislators nationwide, has asserted that, among lawmakers, "opposition to the

law is now coming equally from both political parties."<sup>11</sup>

Nebraska has persuaded federal officials to approve a system that includes portfolios. Washington State allows high school students to take an exit exam up to four times, and, after four failures, students can try alternative means to convince school officials they should be granted a diploma. Utah, a state in which 245 of 810 schools were put on a federal watch list because they failed to make AYP, received a visit in early February 2004 from Ken Meyer, deputy assistant secretary of education, who tried to convince lawmakers and school officials that the state's test was fair. The upshot was that the Republican-controlled Utah House "nevertheless voted 64 to 8 on February 10 not to comply with any provisions not fully financed by federal money."<sup>12</sup>

**3. Do problems remain with the dropout rate and scandals?** Last year, news surfaced about scandals in test administration and in the reporting of scores and dropout rates, particularly in Houston, where it was revealed that dropout rates had been seriously manipulated. After a Houston whistleblower, Robert Kimball, revealed what had gone on in his school, the Texas Education Agency audited the 2000-01 records for the Houston school system and announced that more than half of the students who left the district during that

school year, Secretary Paige's last year as superintendent, had been incorrectly recorded as other than dropping out.<sup>13</sup>

Paige, questioned many times about what happened in Houston, initially declined to answer questions about it. But starting in December 2003, he began to politicize the issue, accusing his critics "of little more than a pre-election effort to undermine the Bush Administration."<sup>14</sup> Secretary Paige has never even admitted that the Texas Education Agency's report, accepted in Texas as accurate, exists. And he continues to defend every aspect of his record in Texas, though always in a very general way.

One major issue in the past year has been the so-called ninth-grade bulge. In January 2004, Boston College researchers released the results of a study conducted for the National Board on Educational Testing and Public Policy in which they concluded that the number of students being held in the ninth grade for failure to pass tests or other academic requirements "has nearly tripled since the late 1960s."<sup>15</sup> In Charlotte-Mecklenburg, North Carolina, "a special watchdog group flagged a 29% bulge among 9th graders as early as 2002."<sup>16</sup> Evidence is growing that once a student is held back the chances of graduation begin to recede. Of course, if a student is held back twice by grade 9, he or she may reach the age at which dropping out of school does not require parental permission.

Two additional problems related to dropouts are bedeviling educators, one statistically clear and the other anecdotal. Dropout rates are stabilizing across the country, but there is a pattern to the rates. According to a study conducted by the Center for Social Organization of Schools at Johns Hopkins University, 2,000 high schools across the nation have dropout rates of 40% or higher, and "those schools are overwhelmingly attended by African American and Latino students."<sup>17</sup> Many of those students leave school in grades 9 or 10 after failing required tests and reaching the legal school-leaving age.

The anecdotal issue involves the pressuring of underperforming students to leave school in ninth grade. This evidence is hard to come by, but it surfaces frequently enough to be worrisome. Walter Haney, a principal author of the Boston College study on the ninth-grade bottleneck, said, "It's just really happening that schools are sacrificing kids to make schools appear to look better. How extensive the phenomenon is we don't know."<sup>18</sup>

#### **4. Has there been improvement in the ability to**

**avoid nonrandom human error and other unanticipated test problems?** In last year's article on testing, I stated, "As long as we rely on a single high-stakes test with a narrow gauge for grade promotion and graduation decisions, we can't get around the problem."<sup>19</sup> New York City provided me with a recent example to buttress that statement.

David Herszenhorn, in a May 2004 article titled "Mishaps Still Plague Citywide Reading Tests," asked readers which of three "mishaps occurred on this year's citywide reading test."<sup>20</sup> Could it be that some teachers — whether inadvertently or purposely — allowed students to practice on last year's exam, which contained some of the questions used on this year's test? Is it possible that some of the questions were displayed before the exam date on a local TV channel? Perhaps it was that the answer sheet from the publisher of the exam, Harcourt Assessment, Inc., did not match the test booklet. City education officials admitted that the answer was "All of the above."<sup>21</sup>

Jane Hirschmann of the group Time Out from Testing stated that some people might wish to see such problems as a quality control issue, but her view is that they underline "why no one should use a single score on a test to determine an 8-year-old's future."<sup>22</sup> In Connecticut in February 2004, the state challenged the test results on a CTB/McGraw-Hill test affecting 120,000 students. Betty Steinberg, the state commissioner of education, was "concerned that recent problems in the scoring of state exams by CTB/McGraw-Hill may be symptomatic of larger problems looming as greater demands are placed on the testing industry." Meanwhile, McGraw-Hill claims to be "working to fix any initial glitches."<sup>23</sup>

Problems with tests are simply endemic as pressures mount to create and score tests faster and faster. In the past five years, there have been problems everywhere from Arizona to Washington and from Michigan to Kentucky. According to David Sadker and Karen Zittleman, "the National Board on Educational Testing and Public Policy reported that 50 high-profile testing mistakes had occurred in 20 states from 1999 to 2002."<sup>24</sup>

Tests are good only if they yield reliable responses. And many external things can get in the way — from fire drills to unexpected noise and other distractions near the test site to student illness or poor weather conditions on the day of the test. Moreover, it is maddeningly difficult to produce tests year after year that are really equivalent. Last spring, Oregon education officials were upset when 82% of the state's 10th-graders

failed a required test of math problem solving, a test that only 50% had failed in the previous year. Since it is not clear what went wrong, “it’s possible the state will statistically boost the test scores to make this year’s result fair to students and schools.”<sup>25</sup>

Cheating remains another issue that interferes with

*In such states as Kansas, Kentucky, Louisiana, Missouri, Nebraska, and South Carolina, plaintiffs have filed lawsuits asking for additional funds for poor schools that cannot afford to provide students with high-quality instruction.*

dependable test results. A *Detroit News* article, based on an ABC News Primetime poll of youngsters aged 12 to 17, reported that “more than seven in 10 teenagers say students in their school cheat on tests.”<sup>26</sup> Test monitors now must worry about students using instant text messaging and e-mail during tests. Judy Miller, speaking from Los Angeles, referred to cheating scandals involving teachers in Milwaukee and Chicago: “The cheating ranges from whispering answers to students during testing to erasing and changing answers.”<sup>27</sup> Miller said that many of the 75 teachers involved in a cheating scandal cited pressure as the reason for their behavior and mentioned the possibility of losing their jobs or being told their schools would be closed.

One of the important features of NCLB is the disaggregation of data by ethnic background, and I cited several examples of the failure to gather accurate data in my article last year. More and more students are the products of more than one ethnic group and so are unsure of what to say; others, particularly young children, simply do not know the answer; and still others, often older students, choose not to respond. I was interested to learn that in 2003 “a record 25% of students declined to disclose their race when registering for the SAT, a nonresponse rate statistically significant enough to call into question ethnic breakdowns of achievement data.”<sup>28</sup>

This year, some unanticipated developments occurred that bear on nonrandom error. Students and parents have become creative about circumventing failing exit exams once children have passed all of the school’s

courses. Florida parents learned that they can send their children’s transcripts to the North Atlantic Regional Schools in Lewiston, Maine, and receive a legal diploma sufficient to have their children admitted to many colleges in spite of a failing high school exit exam. “Their unorthodox maneuver is drawing attention among parents around the country — particularly those with children in special education — who have sought ways around state exit exams they believe are unfairly denying their children diplomas.”<sup>29</sup>

In Wisconsin, the attorney general “has issued an opinion that the federal government can’t force states to comply with the No Child Left Behind law without fully funding it,” a first step in a lawsuit that could end up in the Supreme Court.<sup>30</sup> The

Maryland State Teachers Association and other state education groups have asked the state commissioner of education to halt or slow the imposition of a requirement that students must pass standardized tests to graduate. In such states as Kansas, Kentucky, Louisiana, Missouri, Nebraska, and South Carolina, plaintiffs have filed lawsuits asking for additional funds for poor schools that cannot afford to provide students with the high-quality instruction required to perform well. Steven Smith, an education analyst at the National Conference of State Legislatures, commented, “In recent years, judges have sided with the plaintiffs in about 70% of the adequacy lawsuits, ordering states to come up with extra money. He predicted that the number of lawsuits is likely to rise.”<sup>31</sup>

**5. What role has grade retention played in the past year?** The major news on retention in the past year comes from two of the nation’s largest cities. Chicago and New York arrived at very different decisions on whether students should be retained.

It is clear that retention is not supported by either experience or research. In recent years, Chicago had one of the largest and best-researched retention programs, a program that was sharply curtailed last year after very disappointing results over a nine-year period. Elizabeth Duffrin writes that the “latest studies again found that holding low-achieving students back did not help them academically and increased the likelihood they would drop out.”<sup>32</sup> Melissa Roderick, a director of the Consortium of Chicago School Research, and one of the co-authors of a University of Chicago

research study on retention in Chicago's schools, was crystal clear about what should be done: "They should get rid of retention. It just didn't do anything for these kids."<sup>33</sup>

The National Association of School Psychologists has issued a position paper on retention and social promotion that reveals the uselessness of mass retention. Still, the practice persists:

Despite a century of research that fails to support the efficacy of grade retention, the use of grade retention has increased over the past 25 years. . . . Evidence from research and practice highlights the importance of seeking alternatives that will promote social and cognitive competence of children and enhance educational outcomes.<sup>34</sup>

Surely New York City Mayor Michael Bloomberg and Joel Klein, the chancellor of the city's schools, were well aware of the Chicago experience when retention came up for a vote in March 2004. In fact, one of the members of New York's Panel for Educational Policy is Augusta Souza Kappner, president of Bank Street College of Education. She and a majority of the panel members made clear their intention to vote against grade retention based on the failure of the similar program in Chicago, evidence from dozens of research documents related to the practice, and the fact that retention had failed dismally in New York City in the 1980s under Mayor Edward Koch. Very probably all of the panel members understood that, "over the last two decades, dozens of studies have led many educators to conclude that policies forcing students to repeat a grade are costly and often counterproductive, resulting in little or no improvement in student achievement and sharp increases in dropout rates."<sup>35</sup>

One hates to be cynical about anything connected to education, but the recent actions of New York City's mayor invite more than a dollop of cynicism. In order to get the vote on third-grade retention that he wanted, just hours before the vote, "Mr. Bloomberg fired two of his own appointees to the panel and engineered the dismissal of the member appointed by the Staten Island borough president. Three new appointments were hastily made, and the newly constructed panel held its vote before a stunned and furious audience. . . . It approved Mayor Bloomberg's policy change, 8-5."<sup>36</sup>

But here's the most cynical aspect of the tale. Sometimes, school policy makers and school administrators are accused of holding youngsters back a year before

a major test in order to make themselves look a bit better. And, as Bob Herbert said of Mayor Bloomberg, "Holding back a few thousand additional third graders this year will inevitably lead to higher scores in the benchmark fourth-grade tests next year, which just happens to be a re-election year for Hizzoner."<sup>37</sup>

It's not for lack of knowing what to do that we resort to retention. Smaller classes, individualized tutoring, early intervention, assignment of enthusiastic and well-trained teachers, intelligent use of data, and constant attention frequently yield positive results in high-poverty schools. In 17 of the poorest schools in Montgomery County, Maryland, selected by Superintendent Jerry Weast, just such an approach was started four years ago. "In those schools, where children through the third grade have experienced the programs each year, CTBS [Comprehensive Test of Basic Skills] scores have increased significantly in all subjects since 2000, officials said."<sup>38</sup>

## CONCLUSION AND PREDICTIONS

The efforts of states to craft standards and tests for the past 15 years and the NCLB law are not going away, although I believe considerable reshaping will take place as a result of unrealistic expectations, pressure from a majority of the states, financial problems, and changes in elected and appointed leaders. Thousands of schools that are making Herculean efforts to improve will not stand for being judged failures by some absolute standard tied to an arcane formula. Schools will not be able to attract high-quality teachers to a system that stifles richness and creativity and emphasizes a narrow band of knowledge and a very restricted set of tests to measure that knowledge. Citizens across the country will notice that many private and religious schools do not offer such a narrow curriculum and limiting set of tests, yet their students thrive and gain admission to the best colleges.

Americans will recall that many U.S. schools were actually doing well before the latest national testing movement began. They were graduating young people who went to hundreds of fine colleges, did well, and produced the world's greatest economy and mightiest military force — not to mention high quality in everything from technology and the arts to popular entertainment, sports, and culture. Without question, some middle-class schools were doing a mediocre job, even though they had ample resources, and the shame of the system was that many of the schools in the poorest areas of

our country were neither properly supported nor pressured to make the improvements necessary to produce well-educated graduates. Perhaps schools that demonstrate high achievement for two years in a row should be tested only at intervals of three years. Or perhaps they might register their own assessment measures, which would often be more demanding and revealing than those imposed by the states or the federal government.

The federal government and the states, I predict, will slowly — but not glacially — begin to focus their efforts more properly. There will be a clear movement toward more sophisticated measures that assess a more worthwhile and powerful curriculum and more sensible methods for judging a school's progress. Schools and subgroups cannot remain below some reasonable standard forever, but the fact is that no large city, state, or country — other than mythical Lake Wobegon — has ever produced an entire population of students who are above average. Parents, teachers, and students will begin to assert their influence at least as strongly as elected officials and business leaders, though the latter groups will continue to play an important and welcome role.

Largely because of extensive media coverage, Americans are learning that quite often “objections to testing boil down not to empirical differences but to ideology.”<sup>39</sup> Giving students multiple tests instead of one per subject, reducing the number of subgroups or increasing their minimum size, making the assessment formulas clearer and fairer, taking into increasing account teacher evaluation and various kinds of performance, reducing the number of students retained, and limiting retention to once in a student's K-12 career will all become more common in the future. Narrow ideology must yield to rigorous but attainable standards, multiple forms of assessment, and best practices.

1. Dave Posner, “What's Wrong with Teaching to the Test?,” *Phi Delta Kappan*, June 2004, p. 750.
2. W. James Popham, “Softening Up the ‘No Child’ Act: Is Edge Softening Really Enough?,” *Education Week*, 26 May 2004, p. 40.
3. Linda Perlstein, “School Pushes Reading, Writing Reform,” *Washington Post*, 31 May 2004.
4. Susan B. Neuman, “Reading Experts Offer Insights into State, Federal Policies,” *Education Week*, 12 May 2004, p. 6.
5. Joan L. Herman, Eva L. Baker, and Robert Linn, “Newsletter of the National Center for Research on Evaluation, Standards, and Student Testing,” *CRESSST Line*, Spring 2004, p. 2.
6. Catherine Gewertz, “Debate over Retaining 3rd Graders Roils N.Y.C.,” *Education Week*, 16 June 2004, p. 5.
7. Associated Press, “Study: 9 Out of 10 Schools Won't Meet Federal Standards in a Decade,” *Hartford Courant*, 28 May 2004, p. 1.
8. Michael Winerip, “Making Leaps, but Still Labeled as Failing,” *New*

- York Times*, 28 April 2004, p. A-21.
9. Lynn Olson, “States Seek Federal OK for Revisions,” *Education Week*, 5 May 2004, p. 1.
10. David J. Hoff, “Accountability Conflicts Vex Schools,” *Education Week*, 10 March 2004, p. 23.
11. Michael Dobbs, “More States Are Fighting ‘No Child Left Behind’ Law,” *Washington Post*, 18 February 2004, p. 4.
12. Sam Dillon, “Bush Education Officials Find New Law a Tough Sell to States,” *New York Times*, 22 February 2004, Sunday National Edition, p. A-20.
13. Pam Easton, “Houston Schools Get 6 Months to Improve Records on Dropouts,” *Austin American Statesman*, 9 August 2003, p. B-7.
14. Diana Jean Schemo, “Bush's Education Chief Defends Policy and Past,” *New York Times*, 28 January 2004, p. A-14.
15. Debra Viadero, “Study Probes Enrollment ‘Bulge’ in 9th Grade,” *Education Week*, 28 January 2004, p. 1.
16. *Ibid.*, p. 16.
17. Debra Viadero, “‘Dropout Factories’ Identified in Hopkins Study,” *Education Week*, 14 July 2004, p. 14.
18. Viadero, “Study Probes Enrollment ‘Bulge,’” p. 16.
19. Mark F. Goldberg, “The Test Mess,” *Phi Delta Kappan*, January 2004, p. 363.
20. David M. Herszenhorn, “Mishaps Still Plague Citywide Reading Test,” *New York Times*, 13 May 2004, p. 1.
21. *Ibid.*
22. *Ibid.*, p. 3.
23. Jeff Archer, “Connecticut Tests Delayed by Scoring Glitches,” *Education Week*, 11 February 2004, p. 5.
24. David Sadker and Karen Zittleman, “Test Anxiety: Are Students Failing Tests — Or Are Tests Failing Students?” *Phi Delta Kappan*, June 2004, p. 742.
25. Betsy Hammond, “Sophomores Nose-Dive on Math Problem-Solving Test,” *OregonLive.com*, 21 April 2004, p. 2.
26. *Detroit News*, 30 April 2004, p. 1.
27. Judy Miller, “ABC News on AOL.com/Video@AOL,” 21 May 2004, transcription by Mark Goldberg.
28. “Fewer Students Reveal Ethnicity, Making SAT Achievement Gap Data Unreliable,” *ASCD SmartBrief*, 5 April 2004, p. 1.
29. Sean Cavanagh, “Fla. Students Turn to Maine for Diplomas,” *Education Week*, 28 April 2004, p. 1.
30. Todd Richmond, “Wisc. AG: School Reforms Can't Be Forced,” *boston.com News*, 13 May 2004, p. 1.
31. Michael Dobbs, “Poor Schools Sue for Funding,” *Washington Post*, 6 June 2004, p. 2.
32. Elizabeth Duffrin, “Popular Despite the Research,” *Catalyst Chicago*, 4 May 2004, p. 1.
33. *Ibid.*
34. “Position Statement on Student Grade Retention and Social Promotion,” revision adopted by the Delegate Assembly of the National Association of School Psychologists, 12 April 2003.
35. David M. Herszenhorn, “After Holding Back Third Graders, What to Do?,” *New York Times*, 17 March 2004, p. A-1.
36. Bob Herbert, “Teaching Us a Lesson,” *New York Times*, 19 March 2004, p. A-19.
37. *Ibid.*
38. Linda Perlstein, “Scores at Needy Schools in Montgomery Improve,” *Washington Post*, 24 May 2004, p. 1.
39. Bill Evers and Herbert J. Walberg, “Why Not Put Schools to the Test?,” *Christian Science Monitor*, 12 July 2004, p. 2. ■

Copyright of Phi Delta Kappan is the property of Phi Delta Kappan and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.