## The 16th Bracey Report on

# The Condition of Public Education

When political considerations define the agenda for or shape the findings of research, the results can seem frustrating or silly — but often the losers are students and teachers.

**BY GERALD W. BRACEY**

NEWT GINGRICH has suggested that Democrats run their election campaigns on a simple slogan: "Had enough?"[1] Bush Administration shenanigans and Congressional political ploys so brazen as to be unbelievable — for instance, tying a rise in the minimum wage that would benefit millions to a reduction in the estate tax that would benefit the nation's 7,500 wealthiest families — had me nodding in agreement. Herewith, the year in review.

### DESPERATELY SEEKING STRAWS TO GRASP

"This law is helping us learn about what works in our schools. And clearly, high standards and accountability are working. Over the last five years, our 9-year-olds have made more progress in reading than in the previous 28 combined." So said Margaret Spellings at the No Child Left Behind Summit in April 2006, referring to gains on the National Assessment of Educational Progress (NAEP).[2] In statistical circles, what the secretary is doing is called "cherry picking." And in this instance, careless and self-serving cherry picking, too. Those last five years Spellings spoke of span 1999 to 2004. For two of those years Bill Clinton was President, and it is possible that all that gain — all 7 points — oc-

*GERALD W. BRACEY is an associate for the High/Scope Foundation, Ypsilanti, Mich. His most recent book is* Reading Educational Research: How to Avoid Getting Statistically Snookered *(Heinemann, 2006). He can be reached at gbracey1@verizon.net.*

curred on his watch. In two of those five years, NCLB did not exist. In 2001-02, NCLB (signed into law in January 2002) would have been in existence only three months before a NAEP assessment — had there been one. And given the confusion that reigned from 2002 to 2004 and the hostility between the states and the U.S. Department of Education, it is not likely that much gain occurred then. (Remember that in 2002, then-Secretary Rod Paige accused some states of trying to "ratchet down their standards" and thus of being "enemies of equal justice and equal opportunity and . . . apologists for failure."[3] It's surprising he didn't go on to call them terrorist organizations.)

Spellings' statement is true only if you start in NAEP's first trend year, 1971. Begin in the year of NAEP's previous trend high point, 1980, and the gain would be only 4 points. Then, too, there was no gain from 1999 to 2004 for 13-year-olds and a decline of 3 points for 17-year-olds. And why didn't she mention math trends, since 9-year-olds showed a 9-point gain and 13-year-olds a 5-point gain? Seventeen-year-olds, though, showed a 1-point decline.

Spellings also said, "Scores are at all-time highs for African American and Hispanic students."[4] Well, if she meant reading scores for 9-year-olds, that was true. But it wasn't true for 17-year-old blacks or 13- and 17-year-old Hispanics (13-year-old blacks were at an all-time high by a single point). The statement would have been true, too, in mathematics, except for black and Hispanic 17-year-olds.

The regular NAEP assessment of 2005, though, proved less upbeat. The "regular" NAEP assessments, the ones billed as "the nation's report card," change items over time in conjunction with curricular shifts; the NAEP that yields trends administers the same items at each assessment. In the regular assessment for 2005, fourth-grade reading reached the same level as it had at the onset of NCLB in 2002, and eighth-grade reading declined 2 points. In math, scores rose 3 points for fourth-graders from 2003 and 1 point for eighth-graders.

In reading, the proportion of students at or above the proficient level was static for fourth-graders at 31% and fell for eighth-graders from 33% to 31%. In mathematics, the proportion of fourth-graders at or above proficient rose from 32% to 36%, while for eighth-graders it rose from 29% to 30%. While "the Administration scrambled to put the best face on the numbers and to defend the law that some complain forces a test-driven curriculum on the classroom," Lois Romano reported in the *Washington Post*, Ross Wiener of the Education Trust had a more common reaction: "No one can be satisfied with these results. There's been a discernible slowdown in progress since '03, at a time when we desperately need to accelerate gains. The absence of particularly bad news isn't the same as good news."[5] As for President Bush's comment that the achievement gaps were narrowing and "that's positive and that's important," Wiener countered, "It is meager progress. Students of color and low-income students continue to be educated at levels far below their affluent peers."[6]

For many people, Wiener's views raised the question, "Is NCLB working?" Two reports that appeared within two weeks of each other said "No."[7] Both studies analyzed the regular NAEP assessments, not the trend data.

In the first of these reports, researchers at Policy Analysis for California Education (PACE) at the University of California, Berkeley, examined reading trends for 12 states on both the state tests and NAEP. (The states were Arizona, California, Illinois, Iowa, Kentucky, Massachusetts, Nebraska, New Jersey, North Carolina, Oklahoma, Texas, and Washington.) The states began with much higher proportions of "state-proficient" students; the average gap between state-defined proficiency and NAEP-defined proficiency was 38%; the smallest gap was in Massachusetts at 10%; the largest, in Texas at 55%.

The gap in itself is of no great import. Both state standards and NAEP achievement levels for determining proficiency are wholly arbitrary — both lack any connection to external criteria for validation — and the NAEP levels are far too high. For instance, U.S. fourth-graders were 11th in math and third in science among the 26 nations that participated in the 1995 Third International Mathematics and Science Study (TIMSS). But NAEP found only 18% of fourth-graders to be proficient or better in math and 26% to be proficient or better in science. Still, for analyzing changes over time, rather than for absolute differences, the NAEP levels can be useful. The following table shows the annual gains in fourth-grade reading and math in pre- and post-NCLB years.

| | Reading | | Math | |
|---|---|---|---|---|
| | **State** | **NAEP** | **State** | **NAEP** |
| Pre-NCLB Gains | 2.6 | 0.4 | 2.7 | 1.5 |
| Post-NCLB Gains | 1.9 | -0.2 | 2.9 | 2.4 |

As we can readily see, the post-NCLB gains in reading are smaller on both state tests and NAEP. There is actually a loss on NAEP. In math, the state test gains post-NCLB are about the same as prior to the law, while the gains in NAEP mathematics have picked up. Among the states, only Arkansas managed a reading gain of more than 1% per year on NAEP. In math, the annual gains on NAEP ranged from 1.3% in Illinois to 4.0% in Arkansas. Now, if Arkansas can sustain these gains, it can reach 100% proficiency in math

by 2024 — only a decade late. In Illinois, 100% proficiency in math could be attained by 2057. Both projections are hopelessly optimistic, though, because they are based on the unrealistic assumption that equally large increases in gains will occur each year.

The U.S. Department of Education (ED) took its usual approach to such information: it attempted to defame the messenger. Kevin Sullivan, identified by the *Los Angeles Times* as a spokesman for ED, said that PACE "has a track record of putting out flawed and misleading information about No Child Left Behind."[8]

As this issue went to press around Labor Day, I had no comments from Mr. Sullivan or any ED official on the new PACE study. Nor has anyone at ED addressed the similar study conducted by Jaekyung Lee of SUNY, Buffalo, for the Harvard Civil Rights Project. Given that Douglas Harris, Gene Glass, and Robert Linn reviewed Lee's study, it is not likely to fall victim to ED's derogation. At least, not for its methodology. Although Lee uses quite different methods from those used by the PACE researchers, his study produced results similar to PACE's:

> NCLB did not have a significant impact on improving reading and math achievement across the nation and states. Based on the NAEP results, the national average achievement remains flat in reading and grows at the same pace in math after NCLB than before. In grade 4 math, there was a temporary improvement right after NCLB, but it was followed by a return to the pre-reform growth rate. . . .
>
> NCLB has not helped the nation and states significantly narrow the achievement gap. . . .
>
> NCLB's attempt to scale up the alleged success of states that adopted test-driven accountability policy prior to NCLB (e.g., Florida, North Carolina, Texas) did not work. It neither enhanced [these] states' earlier academic improvement nor transferred the effects of a test-driven accountability system to states that adopted test-based accountability under NCLB. . . .
>
> The higher the stakes of state assessments, the greater the discrepancies between NAEP and state assessment results. These discrepancies were particularly large for poor, black, and Hispanic students.

That last finding — the higher the stakes, the higher the state/NAEP discrepancy — does not describe a perfect relationship, but it also comes as no surprise. Lee and colleague Kenneth Wong of Brown University had earlier constructed an index to measure state accountability levels.[9] Using this index, Lee found that the correlation between the height of the stakes and the size of the NAEP/state discrepancy was +.36. Not huge, but statistically significant.

At the state level, few states showed increases in NAEP reading, and none showed accelerated growth after the law was passed. Lee states that eighth-graders maintained growth similar to that seen pre-NCLB but that fourth-graders showed accelerated growth in math in the post-NCLB years. Lee notes, though, that most of the improved growth occurred between 2000 and 2003 and that growth returned to its pre-NCLB rate afterward. But, as mentioned earlier, NCLB came into existence only in January 2002, and it probably had no influence on what was happening in the last half of the 2001-02 school year. Thus, if NCLB affected math growth in the period from 2000 to 2003, it would have had to work its wonders in a single school year, 2002-03. (There were no NAEP data to examine for 2001-02.) As also mentioned, given the haphazard implementation of NCLB, this seems most unlikely. It is at least possible that the rate increases took place prior to NCLB.

Commenting on the NAEP trend results of 2004, Secretary Spellings declared, "Changing the direction of America's schools is like turning the Queen Mary, a large ship whose captain can't change course on a dime. The goal requires a lot of time and effort, but we are beginning to turn our own Queen Mary around."[10] That might or might not be true of the trend analysis. The analyses of regular NAEP data by PACE and by Lee make it clear that, as far as regular NAEP assessments go, the liner is dead in the water.

## NCLB: A THREAT TO THE NATION'S GLOBAL COMPETITIVENESS?

At that NCLB Summit in April, Spellings also said, "There are certain things you can't teach in a classroom that our students *already have — qualities like creativity, diversity, and entrepreneurship*. Our job is to give them the knowledge and skills to compete. . . . America has always been the most innovative society in the world. And together, we will make sure we always are"[11] (emphasis added).

This might be the first time diversity has been listed with creativity and entrepreneurship as a personal quality, but many people believe that those qualities of creativity and entrepreneurship are what keep the nation competitive in the first place. The minister of education of Singapore certainly thinks so. Tharman Shanmugaratnam told *Newsweek* pundit Fareed Zakaria that Singapore had a test meritocracy while America had a talent meritocracy. "We cannot use tests to measure creativity, ambition, or the willingness of students to question conventional wisdom. These are areas where Singapore must learn from America."[12] Even allowing that the minister is being a bit disingenuous — the last thing a totalitarian society like Singapore wants is a cadre of young people who question conventional wisdom —

his comment on tests rings true.

Zakaria had approached the minister because he was intrigued that kids from Singapore aced tests but that, "ten or twenty years later, it is the American kids who are ahead. Singapore has few truly top-ranked scientists, entrepreneurs, inventors, business executives, or academics. American kids test much worse, but seem to do better later in life and in the real world." A Singaporean father who had lived in the U.S. for a period before returning to his island nation confirmed the minister's assertions, telling Zakaria, "In the American school, when my son would speak up, he was applauded and encouraged. In Singapore, he's seen as pushy and weird." Schooling in Singapore "is a chore. Work hard, memorize, test well." The father placed his son in an American-style private school.

Similarly, Joseph Renzulli, who directs the National Research Center on the Gifted and Talented, housed at the University of Connecticut, had Japanese visitors tell him, "Your schools have produced a continuous flow of inventors, designers, entrepreneurs, and innovative leaders."[13] They noticed American creativity and thought the schools had something to do with it.

It is not only between Asian and American schools that one sees the contrast between passive memorization and active participation in the learning process. A *Washington Post* op-ed from a few years ago described the writer's frustrations trying to get Scottish high-schoolers to discuss Shakespeare. "It took months of badgering before I was able to get my Scottish students to speak up in class. They simply weren't accustomed to asking questions or tossing around their own observations. American schools teach American kids to ask questions. They teach students to be curious, skeptical, even contrary. . . . At their best, they teach kids to challenge the teachers."[14]

But Spellings doesn't get it. She visited a school in 2006 and reported that "the class was full of students asking 'what if' questions."[15] But she doesn't see the connection between questions, creativity, and competitiveness. How does she think American kids got those qualities that "they already have" in the first place? Is it something in the water? She needs a long chat with Robert Sternberg, dean of the College of Arts and Sciences at Tufts University. Sternberg calls creativity a habit. If you don't arrange conditions for people to practice the habit, it won't develop. And, he contends, "the increasingly massive and far-reaching use of conventional standardized tests is one of the most effective, if unintentional, vehicles this country has created for suppressing creativity."[16] There is nothing creative about taking a test. Aside from a few rare exceptions, taking a test is the opposite of asking a question.

As I have said before, we'd better think more than twice about replacing a culture that cultivates asking questions with one that worships high test scores. Somebody needs to give Secretary Spellings a wake-up call.

## TOM FRIEDMAN IS FLAT

My most recent book is titled *Reading Educational Research: How to Avoid Getting Statistically Snookered*. I suggest that readers digest it and then read or reread Thomas Friedman's best seller, *The World Is Flat.* His much-lauded book is a golden treasury of undocumented, carefully chosen, and just plain wrong statistics. At least, that describes the original; I have not thought it worthwhile to acquire the 2006 update. Indeed, the only thing we know for certain from the book is that Tom Friedman visits, dines, and hangs out with lots of rich, important, or famous people.

On page 270, Friedman points out that a remarkable number of "top" math and science students have immigrant parents.[17] "Top" is in quotes because two quite specific events, the 2004 Intel Science Talent Institute (for the finalists of the Intel Science Talent Search) and the 2004 International Math Olympiad, define top. Both involve tiny numbers, which Friedman neglects to mention: 40 individuals in the Intel Institute; 20 in the Math Olympiad. Of these, 24 or 60% in the 2004 Intel Institute and 13 or 65% in the 2004 Math Olympiad were the offspring of immigrants. These are remarkable percentages, but I suggest that generalizing from so small a sample is, er, umm, a bit risky. And that's especially true when one considers that the parents themselves constitute a highly selected group.

These parents did not arrive in steerage. They debarked holding H1-B visas, visas reserved for professional workers. Their first stop in the new land was often a university to obtain an advanced degree. These parents also encouraged their children to pursue math, science, and engineering because they perceived these fields to be well paid, freer of bias, and less apt to look only at applicants who have "connections." Still, they do constitute only .04% of the U.S. population.

Having based a rather large claim on 60 people, Friedman then moved to a conclusion based on a single 18-year-old student's eight-year-old memory of what he studied in the fourth grade. Andrei Munteanu moved here from Romania. Inspired by the movie *Armageddon*, he invented a new algorithm to predict collisions between Earth and asteroids, which earned him a slot as a finalist in the Intel Search. He started school here in seventh grade, which, Friedman declares, "he found a breeze compared to his Romanian school."

Friedman quotes a comment Munteanu made to an *Education Week* reporter: "The math and science classes [cover the same subject matter] I was taking in Romania when I was in fourth grade." While the quotation is correct, "breeze" is Friedman's word. The *Education Week* article says only that Munteanu found "his lessons in Romanian schools noticeably more demanding than those he encountered when he began seventh grade in the United States."[18]

Either way, the quotation leaves the strong impression that Romanian schools offer students accelerated academic trajectories compared to those in U.S. schools. Curiously, although the next paragraphs in *Flat* concern TIMSS results, it apparently never occurred to Friedman to check Munteanu's memory against TIMSS data. It would have spoiled his story. The following data compare U.S. and Romanian students' performance on several TIMSS administrations. (Romania did not participate in any fourth-grade or final-year assessments, so the data come from eighth grade only.)

| | | Math | Science |
|---|---|---|---|
| TIMSS 1995 | Romania | 482 | 486 |
| | U.S. | 500 | 534 |
| TIMSS 1999 | Romania | 472 | 472 |
| | U.S. | 502 | 515 |
| TIMSS 2003 | Romania | 475 | 470 |
| | U.S. | 504 | 527 |

Given the scales used by TIMSS, the U.S. advantage over Romania in mathematics can be characterized as substantial, and the advantage in science as large.[19]

I don't know what kind of education Munteanu received in Romania. I would hazard a guess that some people realized they had a talented kid on their hands and whisked him off to an elite school. (As this issue went to press, Munteanu had not replied to e-mails.) Incidentally, Stuart Anderson of the National Foundation for American Policy, author of the study on these talented students whose data Friedman reported, had a rather different take on what the results meant. Noting that American prosperity has always depended in part on an influx of immigrants, Anderson argues for a "Multiplier Effect": we benefit as a nation from immigrants' talents, but we benefit even more from the talents of immigrants' children. The paper avers that the tightening of H1-B visas after 9/11 was a very bad idea. The proportion of H1-B applications rejected rose from 9.6% in 2001 to 17.8% in 2003.[20]

When he gets around to TIMSS, Friedman arrives at the odd conclusion that it showed "the American labor force to be weaker in science than those of its peer countries." Thirteen-year-olds bubbling in answer sheets do not say much about the quality of the labor force. Friedman admits that U.S. eighth-graders attained higher ranks in 2003 than in 1995, adding, "The worrying news, though, was that the scores of American fourth-graders were stagnant, neither improving nor declining in science or math since 1995. As a result, they slipped in the international rankings as other countries made gains." Friedman quotes Ina Mullis of TIMSS: "Asian countries are setting the pace in advanced science and math." And we know this from fourth- and eighth-grade multiple-choice tests that are aimed at everyone? In fact, in arguing that fourth-graders "slipped" between 1995 and 2003, Friedman is making much ado about very, very little. Table 1 shows the fourth-grade scores from the top to where the U.S. finished.

Of the 15 nations that participated in TIMSS in both 1995 and 2003, two of them, England and Latvia, inserted themselves into the top ranks in math in 2003 after not having been there in 1995. In science, three of them did so. In math, Latvia had finished just below the U.S. in 1995, but England was well down the list. In science, England and Singapore were close to the U.S. score in 1995, but Hong Kong was not. Thus, in each case, only one nation made a substantial gain to overtake the U.S.: England gained 47 points in math; Hong Kong, 34 points in science.

**TABLE 1.**

**Top TIMSS Scorers in Fourth-Grade Math and Science, 1995 and 2003**

| Math | | | | Science | | | |
|---|---|---|---|---|---|---|---|
| 1995 | | 2003 | | 1995 | | 2003 | |
| Singapore | 590 | Singapore | 594 | Japan | 553 | Singapore | 565 |
| Japan | 567 | Hong Kong | 575 | United States | 542 | Japan | 543 |
| Hong Kong | 557 | Japan | 565 | | | Hong Kong | 542 |
| Netherlands | 549 | Netherlands | 540 | | | England | 540 |
| Hungary | 521 | Latvia | 533 | | | United States | 536 |
| United States | 518 | England | 531 | | | | |
| | | Hungary | 529 | | | | |
| | | United States | 518 | | | | |

Friedman doesn't discuss TIMSS gains and losses overall. Too bad. At the eighth-grade level, among the 22 nations with scores from both 1995 and 2003, 13 showed declines in math. Some are large and hard to explain: Sweden, -41; Norway, -37; and Bulgaria, a whopping -51. Of the nine gainers, only three — Latvia (17), Lithuania (30), and Hong Kong (17) — had larger gains than the U.S. (12 points). Collectively, these three nations contain 12.8 million people — almost as many people as metropolitan Los Angeles.

In eighth-grade science, 12 nations showed declines from 1995 to 2003, while the U.S. gained 15 points. The three nations that experienced large gains in math showed them in science as well. The three nations that suffered large losses in math displayed them in science as well.

## CALLING JOHN STOSSEL AND ARMSTRONG WILLIAMS

No, the U.S. Department of Education did not hire these two propagandists to diss the NCES report *Comparing Private and Public Schools Using Hierarchical Linear Modeling*,[21] but it did everything in its power to keep anyone from noticing the report's existence. For one thing, it held onto the report for almost a year. Then it employed a time-honored Washington release strategy: if you don't want people to notice a report, release it on a Friday. The press alert e-mail shows a departure time of 10:58 a.m., but the *New York Times* computer didn't log it in until almost 2:00 p.m. This Friday also happened to be a summer Friday, a Friday when most headlines ran above stories about Israel's war with Hezbollah and Hamas.

When I asked about the lack of coverage, several papers replied that they simply didn't have the time and staff to throw at the story on short notice. Kudos to Diana Jean Schemo and the *New York Times* for getting the story onto Saturday's front page.[22] The study remained relatively unnoticed, though. Only 24 newspapers picked up the Schemo article, and only one, the *San Francisco Chronicle*, was a large-market paper.

There was no press conference, not even a press release with comments from Secretary Spellings. Reg Weaver, president of the NEA, harrumphed that, if the results had shown private schools in a good light, "There would have been press conferences and glowing statements about private schools."

A spokesman for ED, Chad Colby, told *Times* reporter Schemo that he did not expect the study to influence policy. "An overall comparison of the two types of schools is of modest utility," Colby said, emphasizing a caveat that appears in the report as well. Then why not spend ED's limited dollars on something of greater import? No doubt Colby would have perceived the import of the study to be substantially greater had the results supported vouchers. Someone in ED, who insisted on anonymity because of the climate surrounding the report, told Schemo that "researchers were 'extra cautious' in reviewing it and were aware of its 'political sensitivity.' "

Looks to me like ED passed each draft of the report through readability formulas and instructed the authors to make the prose as dense and foggy as possible. With effort, one learns that the study replicated and expanded on an earlier analysis of NAEP math data, conducted by Christopher and Sarah Theule Lubienski (and reported in the April 2006 Research column and in the May 2005 issue of the *Kappan*).[23] The Lubienskis examined only mathematics data because they believed reading to be much more affected by home environments. The NCES study looked at both reading and math.

Conducted by Henry Braun and other ETS researchers, the study first compared public schools to all private schools and then to three different types of private schools: Catholic, Lutheran, and conservative Christian. When the comparisons were made with raw scores, the private schools outperformed the public schools in reading and math in grades 4 and 8. When the scores were adjusted for eth-

nicity, income, parental educational level, teacher certification, student mobility, and student absenteeism (among other variables), the public schools outperformed the private schools at grade 4 and held their own at grade 8. In reading, the public schools scored as well as private schools at grade 4. The private school advantage in reading at the eighth grade fell from 18.1 points to 7.3 points, but it remained statistically significant. Comparisons of public schools to each type of private school did not differ from the overall comparisons, except for eighth-grade mathematics, where conservative Christian schools lagged behind public schools.

On 17 August 2004, the *New York Times* ran a front-page story summarizing an AFT analysis of NAEP charter school data that found that charters did not perform as well as similar public schools.[24] On 25 August 2004, 31 mostly conservative education reformers, led by Paul Peterson and funded by Jeanne Allen's Center for Education Reform, took out a full-page ad in the *Times* criticizing the analysis and the *Times* for running with it.

On 15 July 2006, the *Times* ran its story about the NAEP public/private study. It took Peterson a bit longer to react this time, but, by July 31, he had conducted his own analysis and prepared a 51-page report.[25] It's pretty clear that, as Kevin Franck of People for the American Way put it, "When the Going Gets Tough, Privatization Proponents Get Paul Peterson."[26]

Peterson's principal objections concerned including too many Title I students in the sample and including variables such as absenteeism, number of books in the home, and access to computers at home. He argued that these variables could be *effects* of school choice, not demographic variables outside the control of schools. He developed three models using alternative definitions for the variables in the NCES study. In 11 of 12 comparisons using his models, the private schools remained ahead of the publics after the statistical adjustments.

Peterson claims that his analyses do not prove that private schools are better, only that these kinds of analyses are terrifically sensitive to how variables are defined. My guess is that both the NCES and the Peterson analyses underestimate the achievement of public schools because eligibility for free and reduced-price meals is not a great proxy for poverty. A family making $33,000 a year is still eligible, but I would think that such a family lives in circumstances quite different from a family making half that sum.

## THE SECRETARY STUMBLES

As ED spokesperson Colby said, the NCES public/private study will not affect policy. The Tuesday after the Friday release, Spellings played cheerleader for a new Republican-sponsored voucher proposal. Coming at a time when politicians and the media were both accusing the Bush Administration of focusing on trivial but politically charged issues, such as flag burning and same-sex marriage, people saw this proposal as yet another attempt to "energize the base" of conservatives in time for the fall campaigns. Congressional leaders amplified this perception by saying that the proposal would probably not receive attention before 2007 as part of the reauthorization of NCLB. Asked about the relevance of the legislation in light of the study of public and private schools just described, Secretary Spellings called the results "basically inconclusive," noting that the study had a small sample size.[27]

The timing of the study's release and Spellings' appearance did raise another question: How could an Administration so admired and even envied for orchestrating and choreographing its message stumble so badly? Spellings said that she had learned of the study only by reading the *Times* three days earlier. Prior to that, she didn't know the study existed. Oops! Russ Whitehurst, director of the Institute of Education Sciences at ED, said that he had sent the report to Spellings' office two weeks earlier but had failed to alert Spellings to the report's importance.

## READING FIRST FUMBLES — MAYBE BREAKS THE LAW

"There is no federally prescribed reading program." Thus states the Department of Education at its Reading First home page (www.ed.gov/programs/readingfirst/nclb-reading-first.html). According to analyses by the Success for All (SFA) Foundation, there doesn't need to be.[28] Reading First proposals that don't include the "Michigan List" don't get approved.

Michigan was among the first states to have its Reading First program approved, and it was the first to actually distribute funds. Michigan's proposal did not contain any state review of materials to determine how well various programs reflected principles of scientifically based reading research. It simply listed the five top-selling basals that a group of researchers at the University of Oregon had recommended. The proposal was quickly approved.

Rhode Island, on the other hand, had to submit its Reading First proposal six times. The first two drafts required localities to purchase "high-quality reading programs that meet the test of having a scientific research base." They were rejected because they did "not include the rigorous and clearly defined standards the State will use to evaluate the research base of instructional programs and strategies."

The third draft provided the Michigan List, but it also permitted districts to use other programs if they could justify them in terms of a scientific research base. This draft was rejected as well. The fourth draft was identical to the third, except that the lone sentence allowing districts to adopt their own programs was dropped. The proposal sailed through. (Two other drafts were needed to deal with other issues.) Writes the SFA Foundation, "As soon as Rhode Island had limited its schools to the five basals from the Michigan proposal (which Michigan itself accepted with no scientific review whatever), the reviewers had no further concerns about their reading programs" (p. 12).

The word got around. The SFA Foundation continued, "In multiple reviews of state proposals we obtained for almost every state, there is not a single criticism of any state for restricting Reading First grants to schools using any of the favored basals. In contrast, criticism for states suggesting other programs is constant" (p. 13).

Reading First has also promoted the "three-tier model" approach to beginning reading. This model has no research, scientific or otherwise, to underpin it. "I have been unable to find a single research study that supports this intervention design," said Richard Allington, president of the International Reading Association.[29] One would think that to promote it would then be a violation of the law. Yet, in its (successful) application to provide technical assistance to all 50 states, RMC Corporation "specified that the three-tier model will be the instructional model that will be taught and used across all Reading First Sites."[30] That's *the* model, not *a* model. ED has repeatedly told Success for All "that it should alter its program to fit the three-tier model." Success for All has replied that, to alter SFA as Reading First Director Chris Doherty suggests, would negate 19 years of research on the program's effectiveness. Or, as Allington put it, the three-tier model "seems like a good plan if you wanted to confuse a struggling reader."[31]

The SFA Foundation goes a step further: "Reading First has had an unequivocal effect in increasing the use of commercial programs lacking evidence of effectiveness and reducing the use of nontraditional programs that do have strong evidence of effectiveness — exactly the opposite of what the law requires." Through fiscal year 2006, Reading First spent $5.5 billion.[32]

If the SFA Foundation's analyses seem compelling, those submitted to the Inspector General of ED by the Reading Recovery Council of North American are even more meticulously documented. The council claims that Reading First is trying to wipe it out.[33]

As with SFA, the Reading Recovery Council adduces evidence showing that, while its program has a substantial research base, programs that have little or none are accepted while Reading Recovery is not. Reading Recovery is found in 48 states, but its participation in Reading First is limited to just seven. Reid Lyon, the former "reading czar," after years of calling for randomized field trials to establish the validity of education programs, admitted as much:

> What we originally wanted in Reading First was that if you want to buy a program with federal money, it should have gone through clinical trials to be sure it is effective. But there weren't enough programs that went through that level of rigor; so many programs would be screened out and only a limited number of programs would be available. The Department of Education made the decision to make the criteria more general.[34]

In other words, most programs couldn't meet our scientific standards, so we lowered our standards and accepted programs that told us they were good programs. As Allington observed, "Instead of rigorous research, these advertising materials offer testimonials, cherry-picked case reports, or simple assertions that the product design was influenced by the report of the National Reading Panel."[35]

Reading Recovery is easy to eliminate from Reading First. Because the program depends on one-to-one tutoring, studies showing that such tutoring is no more effective than small-group instruction render Reading Recovery cost-ineffective. But not all studies are equal. In one study used to discredit the program, students were taught by a person who was just starting the yearlong training that Reading Recovery requires; in another, only four children received instruction in Reading Recovery, and the teachers had no training in the program.

The Inspector General's office has not paid a visit to the SFA Foundation in about a year. It has not visited the Reading Recovery Council since it filed its complaint in March 2006.

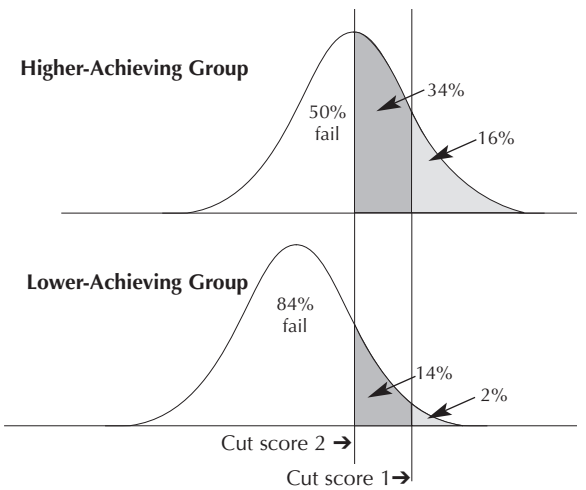
In response to a request from Senate Health, Education, Labor and Pensions Committee Chairman Michael Enzi (R-Wyo.) and ranking Democrat on the committee Edward Kennedy (D-Mass.), the Government Accountability Office announced an investigation of Reading First in October 2005. ED's Inspector General announced "broad audits" of Reading First the following month. Neither agency has issued a report as of September 2006.

## MISLEADING METRICS

My December 2005 Research column warned against using passing rates to measure changes in achievement or

**FIGURE 1.**

**Impact of Cut Scores on Passing-Rate Gap**




Source: Based on La Griffe du Lion, "Closing the Racial Learning Gap," January 2004, www.lagriffedulion.f2s.com/gap.htm.

changes in an achievement gap between different ethnicities. I return to that topic because the practice is more misleading than I had imagined at the time. First, passing rates are arbitrary. Indeed, they might even be arbitrary in the sense of having been negotiated through arbitration. Second, passing rates don't measure how well students perform. They tell you who was able to jump over the barrier you erected; they don't tell you how high the successful jumpers jumped. Worst, they can mask a widening achievement gap. The column presented the hypothetical data below showing that, using passing rates, the black/white achievement gap is closing. But using actual test *scores*, the gap is widening.

**Hypothetical Data**

Score Needed to Pass = 60

| | Pass Rate | Average Score | Gap |
|---|---|---|---|
| Black students, 2004 | 60% | 62 | |
| White students, 2004 | 100% | 78 | 16 |
| Black students, 2005 | 70% | 68 | |
| White students, 2005 | 100% | 92 | 24 |

In fact, the size of any passing-rate gap itself will depend initially on where the cut score is placed. Consider the two curves in Figure 1 and the two different gaps that emerge. Both sets of scores produce bell-shaped distributions, but the lower distribution has a lower mean score. If we set a high cut score (cut score 1), the gap between the percentages passing (shown as shaded areas) appears to be small because few people from either group pass. In the example shown, the gap is 14 percentage points (16% minus 2%). If we use a passing score that fails 50% of the higher-scoring group (cut score 2), it will also fail 84% of the lower-scoring group. Suddenly, our gap has grown from 14 percentage points to 34 percentage points, and all we did was move the cut score. If we were to set a low cut score, the gap would again shrink because most people in both groups would pass.

In the *Wall Street Journal*, Charles Murray, of *The Bell Curve* notoriety, called the use of pass rates "deceptive."[36] Murray started by fingering NCLB as a "disaster for federalism" and contended that "it pushes classrooms toward relentless drilling, not something that inspires able people to become teachers or makes children eager to learn. It holds good students hostage to the performance of the least talented, at a time when the economic future of the country depends more than ever on the performance of the most talented."

Murray then announced his major point: "If you are trying to measure progress in closing group differences, a comparison of changes in pass percentages is inherently misleading." Murray quoted President Bush bragging how Texas had narrowed the black/white gap in pass rates from 35 points to 10 points and said, "President Bush's numbers are accurately stated. They are also meaningless."

Drawing on an analysis by someone who goes by the nom de plume "La Griffe du Lion," Murray showed that, if the mean difference in scores between groups remains constant over time and if one assumes the distribution of scores is normal, then it is possible to predict how changes in the pass rates for one group will affect the size of the passing-rate gap. La Griffe conducted this analysis for two states that some deem models of educational improvement, North Carolina and Texas. The actual pass-rate gaps over time fall almost exactly on the curve that La Griffe predicted.[37]

This is not to say that both North Carolina and Texas did not show improvements overall. They did. I leave it to readers to decide if the states' rising test scores really reflect rising achievement, but rise the test scores did. And various state officials bragged about the closing of the achievement gap. But those officials were looking at reductions in passing-rate gaps.

Do the actual scores converge as the pass rates do? Many consider the Texas tests psychometrically laughable, and the North Carolina Department of Public Instruction advised me that the data "requested are not available."[38]

**TABLE 2.**

**Eighth-Grade NAEP Scores, Texas and North Carolina**

| | Black | White | Gap |
|---|---|---|---|
| **TEXAS** | | | |
| | **Reading** | | |
| 1998 | 246 | 271 | 25 |
| 2002 | 247 | 276 | 29 |
| 2003 | 247 | 272 | 25 |
| 2005 | 246 | 270 | 26 |
| | **Math** | | |
| 1990 | 234 | 272 | 38 |
| 1992 | 243 | 278 | 35 |
| 1996 | 249 | 284 | 35 |
| 2000 | 250 | 286 | 36 |
| 2003 | 260 | 290 | 30 |
| 2005 | 264 | 295 | 31 |
| **NORTH CAROLINA** | | | |
| | **Reading** | | |
| 1998 | 246 | 270 | 24 |
| 2002 | 247 | 274 | 27 |
| 2003 | 247 | 271 | 24 |
| 2005 | 240 | 267 | 27 |
| | **Math** | | |
| 1990 | 231 | 261 | 30 |
| 1992 | 238 | 266 | 28 |
| 1996 | 247 | 277 | 30 |
| 2000 | 252 | 287 | 35 |
| 2003 | 260 | 294 | 34 |
| 2005 | 263 | 292 | 29 |

But eighth-grade NAEP scores, which are available, show very little evidence of any convergence of scores in either state (Table 2).[39]

Thus these results reveal a substantial gain in mathematics in both states, but no gain in reading in either state. In Texas, there is only slight evidence in mathematics of any narrowing of the black/white achievement gap as measured by NAEP scores. What to make of this outcome will depend on whom you ask. I would hold that, for whatever reasons, blacks are not yet receiving effective instruction. Murray and La Griffe would be more likely to hold that the difference is immutable.

Passing rates would be more meaningful if they were connected to anything outside of themselves, but they are not. They are the result of judgments that are sometimes made in highly politicized situations. For instance, the Virginia State Board of Education used a procedure whereby each of 20 members of a committee of judges offered a cut score. Normally, the score chosen would fall near the middle of the distribution of individual cut scores. The board, though, accepted the highest recommended cut score for 25 of 27 tests. For the other two, it set cut scores *even higher* than any of the recommended values. It felt it had to look "tough." It has since had to lower some cut scores because of unacceptably high failure rates.

Readers might recall that an Ohio high school senior received a Golden Apple last year for refusing to take the Ohio exit exam. It was not the act of defiance that garnered the prize, but the reasons given: "I believe these high-stakes tests are biased, irrelevant, and unnecessary." The student then pointed out that Ohio had never taken any action to determine whether they were unbiased, relevant, and necessary. It had done nothing to establish a connection between the tests and the outside world: "In 13 years of testing, Ohio has failed to conduct any studies linking scores on the proficiency test to college acceptance rates, dropout rates, college grades, income levels, incarceration rates, scores on military recruiting tests, or any other similar statistic."[40]

The Commonwealth of Virginia went Ohio one better. Not only did it not conduct any impact study of its high-stakes tests, but when its Technical Advisory Committee consistently called for such a study of consequential validity, the state board dissolved the committee.

In connection with passing rates, George Cunningham, a professor emeritus at the University of Louisville, suggested in an e-mail that the most honest way of constructing a high-stakes test is to decide in advance how many people you want to fail.[41] He's right.

## TWO STATISTICS DEBUNKED

1. *The impact of vouchers.* In 2001, Jay Greene concluded that the rising test scores of Florida schools that had received a grade of F stemmed from the threat of vouchers — a second F grade would make the school's students eligible for vouchers. Wrote Greene:

> This effect of simply having an accountability system in place to put pressure on lower-performing schools operated across all grades, inspiring low-scoring A, B, C, and D schools to improve. But F schools made gains that were even larger than would have been expected simply given how low their previous scores were. The exceptional incentive that existed for schools that had an F grade was the desire to avoid the prospect of school vouchers.[42]

Greene contended that "schools had some motivation to

improve simply to avoid the embarrassment of low scores on the FCAT [Florida Comprehensive Assessment Test]. This motivation operated across all state-assigned grades. But schools with scores of F had a second and very strong incentive to improve to avoid vouchers."

More recently, David Figlio of the University of Florida and Cecilia Rouse of Princeton expressed doubt about the impact of a voucher threat.[43] Looking at changes from 1999 to 2000, Figlio and Rouse found that the F schools gained about twice as many points in reading and math on the high-stakes FCAT as schools with higher grades. They wondered, though, about attributing the gains to the threat of vouchers. Schools that received F's had about 10% higher student mobility than schools that received D's and about 20% higher student mobility than schools receiving higher grades. Controlling for student characteristics could be important.

The analyses Figlio and Rouse conducted also led them to examine gains on a low-stakes norm-referenced test that the state also administered. These gains were much smaller. Further, the gains appear to be limited to the grades in which the high-stakes tests are given. That is, there is a concentration of effort on the skills covered by the high-stakes FCAT, with some spillover to the low-stakes test.

Finally, Figlio and Rouse observed that, prior to implementing the current letter-grade accountability system, Florida placed low-scoring schools on a "critically low-performing" list. At the time, such schools suffered no threat of vouchers; the only incentive to improve would have been to remove the stigma of being on the low-performing list. If the voucher threat is causing the gain, then schools given an F in the new system, where vouchers are imminent, should gain more than schools in years past that had been placed on the critically low-performing list. But Figlio and Rouse found the subsequent performance of these schools to be the same as the performance of schools that more recently faced the threat of getting an F and losing their students to private schools.

In addition, they found no evidence that schools receiving an F that had previously been on the "critically low-performing" list reacted any differently from those receiving an F that had not been on that list. This they take as evidence that the stigma of an F, not the threat of a voucher, motivates improvement.

Figlio and Rouse also think that, because the gains are limited to the high-stakes grades, Florida should test in more grades. That, of course, assumes that rising test scores equal rising achievement, something many of us would not be willing to assume, especially under the current distorting circumstances.[44]

2. *The rarity of high-flying schools.* When I give a presentation, I usually have with me an overhead showing how many "high-flying" schools there are in the state in which I am speaking, plotted against the poverty level of the schools. This is information I take from the interactive database that the Education Trust used for *Dispelling the Myth* (DTM).[45] The Education Trust claimed to have found 3,592 high-flying schools. High-flyers were schools with either 50% or more of their students in poverty or 50% or more of their students being minorities or both, but which nonetheless managed to score in the upper third of schools on the state's test.

Compared to the number of high-poverty schools, the number of high-flying schools is small. But the Education Trust and others have implied that it is a large number and implied more strongly that, if these schools can do it, any schools can. *Dispelling the Myth*, along with *No Excuses* — a miserable report from the Heritage Foundation[46] — and more recently NCLB all imply that schools alone are responsible for the inequitable outcomes of education. NCLB demands that schools — and schools alone — close the achievement gap. NCLB thus forces dichotomous explanations of low performance: you either blame the kids or you blame the schools. NCLB and others blame the schools. Bill Evers of the Hoover Institution said, "It's not just a few, rare schools that succeed, it's thousands of schools. . . . We'd better not hear that racist nonsense anymore."[47]

If one examines the Education Trust's criteria for the high-flyer designation more closely, its argument falls apart. To

attain this elevated status, a school needs only one grade to perform well in one subject in a single year. Thus, if the fourth grade at a school scores in the upper third in reading in 2006, the whole school gets designated as a high flyer, even if the kids can't add whole numbers and the other grades are all below average.

Douglas Harris of Florida State University examined the data, systematically varying the criteria for determining high-flyer status.[48] First, he simply replicated the Education Trust results using a somewhat more extensive database, ED's School-Level Achievement Data Base, which contains information on 62,000 U.S. schools. Of the 21,234 high-poverty schools (identified using the Education Trust's definition), only 16% qualified as high-flying. Of the 12,869 schools that were both high-poverty and high-minority, only 10% made high-flying status. And that's using the Education Trust's extremely lenient criteria.

But if a school is truly high performing, it ought to show its colors in more than one subject, in more than one grade, and for more than one year. Table 3 shows the percentages of schools that would be identified as "high performing" as these three criteria are progressively increased to the point that the designation requires two years, two subjects, and two grades.

Row 1 (the 1-1-1 row) gives the information for schools that made it to high-performing status in *either of the two* years, *either of the two* subjects, and *either of the two* grades. Each row stiffens the criteria until Row 8, which shows the percentages of schools that made it for both years, both subjects, and both grades. Row 9 shows the percentage of schools that made it under the criteria used by the Education Trust — one year, one subject, one grade.

With criteria as lenient as those in Row 1, most low-poverty schools show up as high-performing schools, while even 30.5% of high-poverty schools and 22% of high-poverty, high-minority schools make the grade. As more years, subjects, or grades are added, though, the proportion of high-flyers diminishes until in Row 8 only 1.1% of high-poverty schools and 0.3% of high-poverty, high-minority schools are designated high-flyers. Twenty-two times as many low-poverty schools as high-poverty schools attain high-flyer status. The likelihood that a low-poverty, low-minority school will attain high-performing status is 89 times greater than for a high-poverty, high-minority school.

## THE LONG LIVES OF STATISTICS

Over the years, I've named laws for other people, but now I'll risk naming my own: Bracey's Law of Statistical Longevity. The law says that any statistic or statistics-based contention that reflects badly on American public schools, whether true or not, will enjoy a long life. For instance, George Will's mutant 1993 creation — "Nationally, about half of urban public school teachers with school-age children send their children to private schools" — lives on and usually appears with the word "urban" deleted. The real figure is 21.5% for urban teachers, 10.6% for all teachers (compared to 12.2% for all families).

Or consider the contention that in 2004 China produced 600,000 engineers, while India generated 350,000. The United States turned out a puny 70,000. These numbers first appeared together in *Fortune* in the summer of 2005 and soon showed up in many newspapers and in a press release from the National Academies.[49]

**TABLE 3.**

**Percentages of Schools Designated High Performing Based on Varying Criteria**

| Row | Criteria: Years | Criteria: Subjects | Criteria: Grades | High-poverty high-performing schools | Low-poverty high-performing schools | High-poverty, high-minority high-performing schools | Low-poverty, low-minority high-performing schools |
|---|---|---|---|---|---|---|---|
| **1** | 1 | 1 | 1 | 30.5 | 80.0 | 22.0 | 84.0 |
| **2** | 2 | 1 | 1 | 12.9 | 59.1 | 7.5 | 63.5 |
| **3** | 1 | 2 | 1 | 14.7 | 62.3 | 9.1 | 66.8 |
| **4** | 1 | 1 | 2 | 11.0 | 56.5 | 6.4 | 60.9 |
| **5** | 2 | 2 | 1 | 4.5 | 41.0 | 2.0 | 44.8 |
| **6** | 2 | 1 | 2 | 3.6 | 37.9 | 1.4 | 41.4 |
| **7** | 1 | 2 | 2 | 2.4 | 33.2 | 0.9 | 36.4 |
| **8** | 2 | 2 | 2 | 1.1 | 24.2 | 0.3 | 26.7 |
| **9** | Education Trust Definition* | | | 15.6 | 54.2 | 10.4 | 56.7 |

*Remember that the Education Trust definition requires high performance in the same year for a single subject and grade.

Some expressed skepticism about the figures from the beginning. Carl Bialik, who writes a column for the *Wall Street Journal* called "The Numbers Guy," located several experts who had doubts, but none could pin down an exact figure.[50]

That was left to Gary Gereffi, Vivek Wadhwa, and a team of researchers at Duke University. Calling this a "mangoes to litchis" comparison, they determined that the actual numbers were 351,537 for China, 112,000 for India, and 137,437 for the U.S.[51] Many who carry the title of "engineer" in India and China would be called "technicians" in the U.S. (I dealt with this report at length in the April 2006 Research column.) Still, months after the Duke study appeared, the larger statistics prevailed in speeches by Secretary of Education Spellings, by Secretary of Commerce Carlos Gutierrez, and by Sen. John Warner (R-Va.).

In May, I published "Heard the One About the 600,000 Chinese Engineers?" in the *Washington Post*.[52] A number of out-of-work engineers sent e-mails pointing to a *job shortage* and indicating that many people holding engineering degrees seek employment in other fields because the other fields pay better. The *Post* piece turned up on hundreds of websites, none of which were apparently visited by journalists Fareed Zakaria and Hedrick Smith or by Alan Simone, president of the Rochester Institute of Technology, who continued to parrot the 600,000 figure. There is no "engineering gap," but no doubt the hordes of Chinese engineers will live on statistically for many years.

In his 2006 update of *The World Is Flat*, Friedman took the Duke study into account and reported its essential findings. But he said he would bet that many of the engineering degrees issued by American universities went to foreign students who would return home. He also expressed "no doubt" that the quality of engineering studies in India and China would catch up to that of the U.S., invoking the example of the recent trend in basketball. In reaction, Wadhwa published an article in *Business Week*[53] that laid out the following salient facts:

• There is no shortage of engineers in this country. Salaries for engineers have dropped — something that wouldn't have happened under conditions of increasing demand. In fact, says Wadhwa, raise the salary level of engineers to that of doctors and lawyers and people will be tripping over one another trying to get into engineering school.

• Between 25% and 40% of engineering graduates don't become engineers. Often, they choose investment banking, real estate, and management consulting as substitute careers.

• A surplus of engineers would create unemployment.

• We've got enough qualified computer programmers. Bill Gates has said that American high school graduates are too unskilled to work for Microsoft, but the company gets about 60,000 applications a month for 2,000 openings. (If Microsoft is having trouble finding employees, it might be because the company has never paid top dollar. It compensated for this with stock options, which worked well in the years of rapid expansion but not so well these days.) Still, Gates and Intel CEO Craig Barrett routinely call for more engineers, and a House committee recommended pumping up the number of engineering graduates by 100,000.

• Most engineering undergraduates are not foreign nationals. The American Society for Engineering Education reports that for the last seven years 92% of undergraduate degrees in engineering went to U.S. citizens or people with permanent residence.

• More than 57% of graduate engineering degrees are awarded to U.S. students, down from 60.3% in 1999. Harvard economist Richard Freeman says this is because salaries for scientists and engineers are lower than for other professions and the benefits of graduate degrees are not worth the costs.

• The majority of foreign-born engineers stay here. The National Science Foundation found that the percentage of stayers rose from 49% in 1989 to 71% in 2003. Wadhwa expects that percentage to decline but says he's willing to bet Friedman that it won't fall to 1989 levels.

Then there's the statistics-based claim that money doesn't matter, first uttered by Eric Hanushek in 1989: "There is no strong or systematic relationship between money and achievement."[54] As Keith Baker was quick to point out in the April 1991 *Kappan*, Hanushek established no criteria for what a "strong or systematic relationship" would look like, used a primitive method that prevented any such determination, and actually found that the number of positive out-

comes was vastly larger than the number of negative.[55] That indicates that, under some circumstances, money might very well have a strong and systematic relationship to achievement. In the ensuing years, the issue was hotly debated, especially with regard to the efficacy of additional money to reduce class size. Now, in the summer edition of *The American Enterprise Online*, comes Jay Greene to write as if Hanushek's research were definitive and no debate ever existed: "Economist Eric Hanushek of Stanford University examined every solid study on spending and outcomes — a total of 163 research papers — and concluded that extra resources are more likely to be squandered than to have a productive effect."[56]

There are a number of things to be said about this statement. First, there were not 163 research papers, only 91, and many of those had no bearing on expenditures. Hanushek was also looking at the impact of teacher salaries, teacher experience, teacher education, and teacher/pupil ratio. There were 163 *estimates* of the impact of money on achievement. Hanushek counted every estimate from every comparison. That is, if a study estimated the impact for elementary schools, it produced one estimate. If it analyzed for grades 1-6 separately, it produced six estimates. If it analyzed for grades 1-6, gender, and the four largest minority groups, it produced 48 estimates. When Princeton economist Alan Krueger analyzed the Hanushek data on class size, he noted that nine studies produced 122 estimates and that just two studies alone generated 48. When he opted to count each study only once, Krueger reversed Hanushek's conclusion.

Second, many of the studies were not "solid" (whatever that means). Thirteen percent of them failed to have any sign — positive or negative — attached to their results, a sure indication that they would not pass anyone's muster for "scientifically based research." The two studies that generated the 48 estimates were both by the same authors and carried the titles "The Merits of a Longer School Day" and "Classmates' Effects on Black Student Achievement in Public School Classrooms." The entire list of titles can be found in the Summer 1997 issue of *Educational Evaluation and Policy Analysis*.

Third, Hanushek's analysis bore on the *level* of achievement, but his recommendations applied to *changes* in achievement. Parental levels of education, family income, and so on greatly affect level of achievement. These demographic variables have less impact on changes in achievement.

Finally, Hanushek never used the word "squandered," nor did he imply it through innuendo.

The "Big Lie" propaganda technique initially meant telling a whopper so big that common folk, themselves accustomed to telling only small lies, would believe the big lie because they could not imagine that anyone would have the audacity to bend the truth so much. Later, the notion that if you tell a lie often enough people will come to assume it's true was added. And so it goes with the "money doesn't matter" people.

Finally, there is the curious case of the mutant second-grade test-score statistics. My website (at www.americatomorrow.com/bracey) received a query from someone in Buffalo about how states use the number of kids who read below grade level in second grade to project future prison construction needs. It seemed a well-established belief in the African American community.

The first reaction among members of my Education Disinformation Detection and Reporting Agency (EDDRA) was that this was an "urban legend," especially because Google searches turned up a variety of statements and grades. But the searches also turned up three legitimate, albeit secondary, sources. A 2004 *Washington Post* story by Andrew Block of the Just Children Program in Charlottesville, Virginia, and Virginia Weisz of the Children's Rights Project in Los Angeles contained this: "In California, correctional officials reportedly look to the percentage of children who never make it past fourth-grade reading level to help them gauge the number of future prison beds to fund."[57]

That immediately struck me as improbable. Knowing how many kids "never make it past fourth-grade reading level" would be a logistical and record-keeping nightmare. I decided the operative word was "reportedly," and when I contacted Block, he agreed, saying the Corrections Department in California had vehemently denied the practice.[58]

Consultant Mike Schmoker had written in a 1999 *Education Week* article that the state of Indiana found it useful to base projections for future prison construction on the number of second-graders who weren't reading on grade level.[59] Similarly, Linda Katz of the Children's Literacy Initiative had written, "Indiana's former governor has stated that determining the number of new prisons to build is based, in part, on the number of second-graders not reading at second-grade level."[60]

The "number of second-graders" is an example of "mutant statistics" — statistics that begin as legitimate numbers but then get transformed into something false. As I noted above, George Will's statement that about half of teachers in urban public schools send their kids to private schools is one such. Will had originally written that about half of Chicago public school teachers send their children to private schools. This was true, although Will failed to mention that religious considerations played a large role

in that decision. But after stewing in Will's brain juices for six months, the statistic mutated into a national number.

The source of the test scores/prison building statistic that mutated proved to be Sen. Evan Bayh (D-Ind.), formerly governor of Indiana. In a book and in articles and speeches, Sen. Bayh expressed his belief in the need for early intervention to keep kids out of jail later: "I remember meeting with Jim [Corrections Chief Jim Aiken] one afternoon in my office. I asked him to explain to me how he could predict the number of criminals we'd be incarcerating in the future. 'We've got this equation,' he said. 'And it's got a lot of variables in it. But the single most reliable predictor is the number of at-risk children in second grade today.'"[61] One presumes that by "reliable" Aiken probably meant "powerful."

Bayh's interest was in destroying the power of that statistic: "In other words, we look at the circumstances currently facing *eight-year-olds* in order to gauge how full our jails will be six or seven years down the road. If ever there was a powerful argument for early intervention, for ensuring that kids grow up in the best possible circumstances, this is it." He went on to talk about those circumstances, especially the importance of having a father around.

When I first started writing these reports, neither I nor the *Kappan* editors were certain that sufficient data would arrive each year to justify an annual publication. Each year, though, an increasing abundance of information has presented a different problem: with a relatively fixed amount of space, what to cut? George Kaplan's 1982 observation that education's story doesn't break, it oozes, is no longer true.

To cope with all the information, we've decided to place a number of additional segments on my website, EDDRA. If readers just want to peruse the reports or the Rotten Apples in Education Awards for past years, they can do so at the EDDRA archive, www.america-tomorrow.com/bracey/EDDRA. If readers want to receive occasional bulletins from me and other EDDRA members (not usually archived), they'll need to fill out the form found at www.america-tomorrow.com/bracey. Membership is free.

If you visit the EDDRA site, expect segments on Harvard or Bust, the desperate quest for a "brand" university; Whatever It Takes, programs to reconnect youths to schools; Errors by Testing Companies; The Constricting Curriculum; What Doesn't Work in the What Works Clearinghouse; some updates on topics covered here; and, of course, the Golden Apples in Education Awards for 2006.

A number of topics appropriate to this report have been dealt with over the year in the Research column: the state of literacy (May 2006), dropouts (June 2006), and the disappearance of kindergarten and recess (November 2005 and January 2006).

I close by noting three things. First, at a luncheon to launch the Committee for Economic Development's campaign for universal preschool, a speaker emphasized the importance to businesses of workers around the age of 40. Second, the World Economic Forum ranks the U.S. "number one" in global competitiveness.[62] Third, the quintessential "mediocrities" of this nation were supposed to have been the members of the high school class of 1983, those who left school a mere six weeks after *A Nation at Risk* warned of a "rising tide of mediocrity." In 2006, the class of 1983 turned 40. Might there be a connection?

---

1. Quoted in Karen Tumulty and Mike Allen, "Republicans on the Run," *Time*, 3 April 2006.
2. "Remarks by Secretary Spellings at No Child Left Behind Summit," U.S. Department of Education, 27 April 2006, www.ed.gov/news. Search on "Press Releases."
3. "Letter Released from U.S. Education Secretary Paige to State School Chiefs on Implementing No Child Left Behind Act," U.S. Department of Education, 23 October 2002.
4. "Remarks by Secretary Spellings."
5. Lois Romano, "Test Scores Move Little in Math, Reading," *Washington Post*, 20 October 2005, p. A-3.
6. Ibid.
7. Bruce Fuller et al., *Is the No Child Left Behind Act Working?* (Berkeley: Policy Analysis for California Education, 2006); and Jaekyung Lee, *Tracking Achievement Gaps and Assessing the Impact of NCLB on the Gaps* (Cambridge, Mass.: Civil Rights Project, Harvard University, June 2006).
8. Mitchell Landsberg, "Reading Gains Slowing, Study Says," *Los Angeles Times*, 20 June 2006.
9. Jaekyung Lee and Kenneth Wong, "The Impact of Accountability on Racial and Socioeconomic Equity: Considering Both School Resources and Achievement Outcomes," *American Educational Research Journal*, Winter 2004, pp. 797-832.

10. "Spellings Hails New National Report Card Results," U.S. Department of Education, 14 July 2005, www.ed.gov/news. Search on "Press Releases."

11. "Remarks by Secretary Spellings."

12. Fareed Zakaria, "We All Have a Lot to Learn," *Newsweek*, 9 January 2006, www.msnbc.msn.com/id/10663340/site/newsweek.

13. Joseph Renzulli, "A Quiet Crisis Is Clouding the Future of R & D," *Education Week*, 25 May 2005, pp. 32-33.

14. Amy Biancolli, "At Least Our Kids Ask Questions," *Washington Post*, 27 April 2001, p. A-23.

15. "Remarks by Secretary Spellings."

16. Robert J. Sternberg, "Creativity Is a Habit," *Education Week*, 22 February 2006, p. 47.

17. Thomas L. Friedman, *The World Is Flat* (New York: Farrar, Straus and Giroux, 2005).

18. Sean Cavanaugh, "Immigrants' Children Inhabit Top Ranks of Math, Science Meets," *Education Week*, 28 July 2004, p. 14.

19. Albert E. Beaton et al., *Mathematics Achievement in the Middle School Years* (Chestnut Hill, Mass.: Boston College, 1996); Albert E. Beaton et al., *Science Achievement in the Middle School Years* (Chestnut Hill, Mass.: Boston College, 1996); Ina V. S. Mullis et al., *TIMSS 1999 International Mathematics Report* (Chestnut Hill, Mass.: Boston College, 2000); Michael O. Martin et al., *TIMSS 1999 International Science Report* (Chestnut Hill, Mass.: Boston College, 2000); and Patrick Gonzales et al., *Highlights from the Trends in Mathematics and Science Study: TIMSS 2003* (Washington, D.C.: National Center for Education Statistics, 2005, Report No. 2005005).

20. Stuart Anderson, "The Multiplier Effect," *International Educator*, Summer 2004, pp. 14-21, available at www.nfap.com.

21. Henry Braun, Frank Jenkins, and Wendy Grigg, *Comparing Private Schools and Public Schools Using Hierarchical Linear Modeling* (Washington, D.C.: National Center for Education Statistics, U.S. Department of Education, July 2006, Report No. NCES 20060461).

22. Diana Jean Schemo, "Public Schools Perform Near Private Ones in Study," *New York Times*, 15 July 2006, p. A-1.

23. Christopher Lubienski and Sarah Theule Lubienski, *Charter, Private, Public Schools and Academic Achievement: New Evidence from NAEP Mathematics Data* (New York: National Center for the Study of the Privatization of Education, Columbia University, 2006).

24. Diana Jean Schemo, "Nation's Charter Schools Lagging Behind, Test Scores Reveal," *New York Times*, 17 August 2004.

25. Paul E. Peterson and Elena Llaudet, "On the Public-Private School Achievement Debate," Program on Education Policy and Government, August 2006, www.ksg.harvard.edu/pepg/research.htm.

26. Kevin Franck, "Cutting Through Right-Wing Spin on Public Education: When the Going Gets Tough, Privatization Proponents Get Paul Peterson," www/educationnews.org/Commentaries/Cutting_Through_Right-Wing_Spin.htm.

27. Diana Jean Schemo, "Republicans Propose National School Voucher Program," *New York Times*, 19 July 2006, p. A-17.

28. Success for All Foundation, "The Mismanagement of Reading First: Summary of Evidence, Part 1," http://susanohanian.org/show_research.html?id=128; and idem, "The Mismanagement of Reading First: Summary of Evidence, Part 2," www.susanohanian.org/show_research.html?id=129.

29. Richard Allington, "Research and the Three Tier Model," *Reading Today*, April 2006, p. 20.

30. Success for all Foundation, "The Mismanagement," Part 2.

31. Allington, op. cit.

32. "Reading First: Over $4.3 Billion to Improve the Reading Skills of Young Children," U.S. Department of Education, 3 August 2005, www.ed.gov/programs/readingfirst/nclb-reading-first.html. Adding fiscal year 2006 brings the total to $5.5 billion.

33. "Evidence Ignored, Learning Denied: The Attack on Reading Recovery," Reading Recovery Council, March 2006, www.readingrecovery.org/sections/home/oigrequest.asp.

34. Nancy Salvato, "Effective Reading Programs Share Common Characteristics: An Interview with G. Reid Lyon," www.educationnews.org, 18 January 2006.

35. Richard Allington, "What Counts as Evidence in Evidence-Based Education?," *Reading Today*, December 2005, p. 16.

36. Charles Murray, "Acid Tests," *Wall Street Journal*, 25 July 2006, p. A-12.

37. La Griffe du Lion, "Closing the Racial Learning Gap," January 2004, www.lagriffedulion.f2s.com/gap.htm.

38. David Bryant, personal communication, e-mail, 3 August 2006.

39. To access the U.S. Department of Education's NAEP data, go to http://nces.ed.gov/nationsreportcard and click on "Analyze Data."

40. John Wood, "Why I Didn't Graduate," *Education Week*, 25 June 2005, p. 46.

41. George K. Cunningham, personal communication, e-mail, 17 April 2006.

42. Jay P. Greene, "An Evaluation of the Florida A-Plus Accountability and School Choice Program," Manhattan Institute, 2001, www.manhattan-institute.org/html/cr_aplus.htm.

43. David N. Figlio and Cecilia Elena Rouse, "Do Accountability and Voucher Threats Improve Low-Performing Schools?," *Journal of Public Economics*, January 2006, pp. 238-55.

44. See Sharon Nichols and David Berliner, "The Inevitable Corruption of Indicators and Educators Through High-Stakes Testing," Education Policy Studies Laboratory, Arizona State University, Tempe, 2005.

45. Craig Jerald, *Dispelling the Myth: Revisited* (Washington, D.C.: Education Trust, 2001). To access the interactive database, go to www.edtrust.org, click on "Data Tools and Presentations" and then on "DTM Online."

46. Samuel Casey Carter, *No Excuses: Lessons from 21 High-Performing, High-Poverty Schools* (Washington, D.C.: Heritage Foundation, 2000).

47. William Evers, "No More Excuses," *NW Brainstorm*, February 2002, pp. 22-24.

48. Douglas N. Harris, *Ending the Blame Game on Educational Equity: A Study of "High Flying" Schools and NCLB* (East Lansing, Mich.: Great Lakes Center for Education Research and Practice, 2006).

49. Geoffrey Colvin, "America Isn't Ready: Here's What to Do About It," *Fortune*, 25 July 2005, pp. 70-82.

50. Carl Bialik, "Sounding the Alarm with a Fuzzy Stat," *Wall Street Journal*, 27 October 2005. Unlike most of the *Journal*, Bialik's often interesting columns can be obtained free at www.WSJ.com/NumbersGuy.

51. Gary Gereffi and Vivek Wadhwa, "Framing the Engineering Outsourcing Debate: Placing the United States on a Level Playing Field with China and India," Duke University, Durham, North Carolina, 2005.

52. Gerald W. Bracey, "Heard the One About the 600,000 Chinese Engineers?," *Washington Post*, 21 May 2006, p. C-3.

53. Vivek Wadhwa, "Engineering Gap? Fact and Fiction," *Business Week*, 10 July 2006, available at http://businessweek.com/print/smallbiz/content/jul2006/sb20060710_949835.htm.

54. Eric Hanushek, "The Impact of Differential Spending on School Performance," *Educational Researcher*, May 1989, pp. 45-51.

55. Keith Baker, "Yes, Throw Money at the Schools," *Phi Delta Kappan*, April 1991, pp. 628-30.

56. Jay Greene, "Education Myths," *American Enterprise Online*, 2006, www.taemag.com/issues/articleID.19233/article_detail.asp.

57. Andrew Block and Virginia Weisz, "Choosing Prisoners over Pupils," *Washington Post*, 6 July 2004, p. A-19.

58. Andrew Block, personal communication, e-mail, 17 July 2006.

59. Mike Schmoker, "The Quiet Revolution in Achievement," *Education Week*, 3 November 1999, p. 32.

60. Linda Katz, "The Importance of Investing in Literacy," 2000, www.clientheweb.org/investing1.html.

61. Evan Bayh, *Father to Son: A Private Life in the Public Eye* (Carmel, Ind.: Guild Press of Indiana, 2003).

62. *Global Competitiveness Report, 2005-2006* (Geneva, Switzerland: World Economic Forum, 2005). Some reports had Finland in first place. Finland was first in what the WEF calls "growth competitiveness," with the U.S. second. On overall competitiveness, though, the U.S. is first, Finland second (p. 28).