

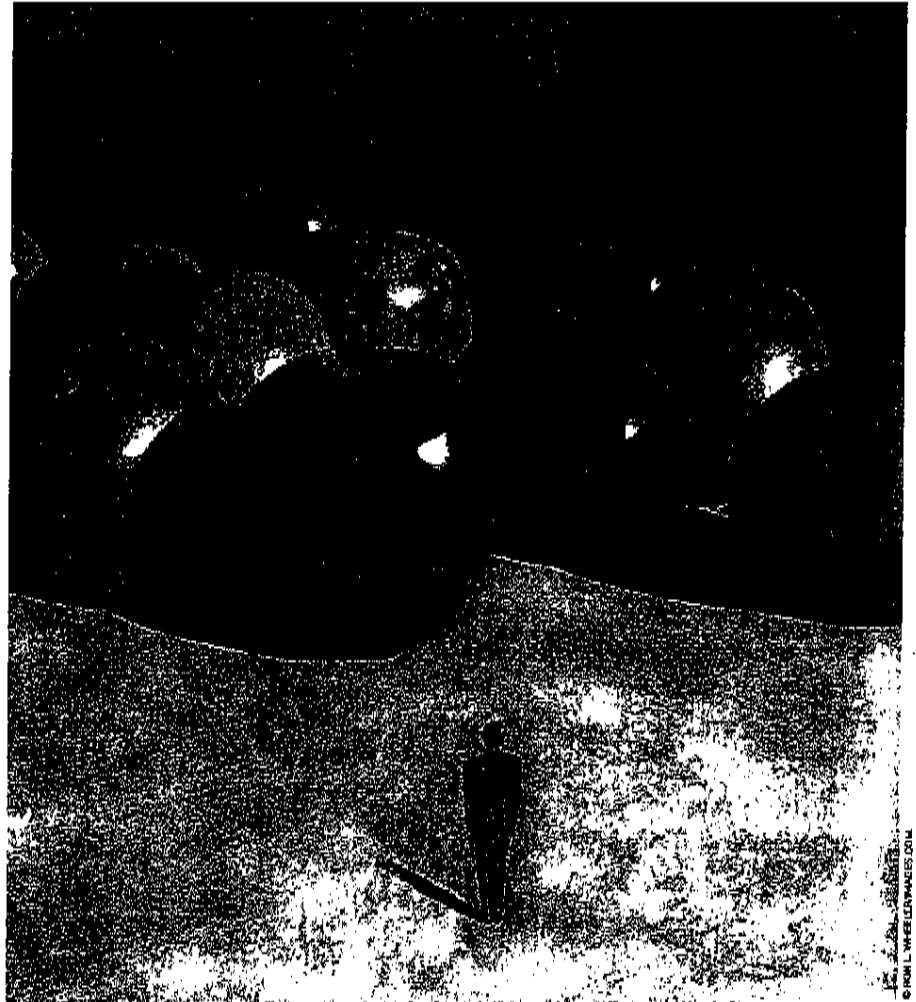
How to Avoid

Education statistics are rarely neutral. These principles can help educators interpret what the research really says.

Gerald W. Bracey

There are three kinds of lies: lies, damned lies, and statistics." This quote from Benjamin Disraeli, Prime Minister of England under Queen Victoria, demonstrates that the field of statistics has needed to defend its honor since its inception in Europe centuries ago.

The original term for statistics, *political arithmetic* (Best, 2001), might be more accurate. Statistics are rarely neutral. Those who collect them have a purpose—sometimes benign, sometimes not—and translate the information to serve that purpose. For example, some people, including representatives of the pharmaceutical industry, say that statistics reveal an "obesity crisis" in the United States. Other people, including some financed by the food industry, allege that the "obesity crisis" is a false alarm spread by drug companies that want the standards for diseases constantly made stricter so that they can define more people as patients and sell them expensive drugs. It's true that the numbers accepted as indications of high chole-



terol, high blood pressure, and high blood sugar have all become much lower in the last decade.

Contradictory claims like these may be one reason why people say that you can prove anything with statistics. You can't, but people will certainly try to prove their particular viewpoints by using only those numbers that serve their purposes.

More Than One Number

You need more than one statistic to get a full picture of just about anything. Educators whose performance is being judged solely by annual standardized test scores will appreciate this point.

As I was writing this, an article in *The New York Times* gave three statistics for various nations' greenhouse gas emissions: total emissions, per capita

Statistical Traps

emissions, and emissions per industrial output (Barringer, 2006). Using total emissions, the United States is number one by far, with China second and Russia third. Using per capita emissions, the United States is still number one and Russia is still third, but Canada is second. (China has lots of people and is still largely a rural nation despite its

about \$7 for every dollar invested in these programs (Berrueta-Clement, Schweinhart, Barnett, Epstein, & Weikart, 1984; Reynolds, 2001).

Principles of Data Interpretation

Despite the limitations of individual statistics and public cynicism about being able to prove anything, people

sioners listed 13 indicators of the "risk," all referring to test scores. For example, "Over half the population of gifted students do not match their tested ability with comparable achievement in school." Given that achievement tests at the time were the principal means of selecting kids for gifted and talented programs, how could this possibly be true? When I sought an answer from some commissioners and their staff members, no one could remember where this statistic came from. How convenient.

Another statistic was, "Average tested achievement of students graduating from college is also lower." The United States has no program to test students graduating from college that would yield a statistic showing their "average tested achievement." What on earth could this mean? These examples illustrate a vital principle of data interpretation: If you find a statement the least bit suspect, ask to see the raw data.

Beware of Selectivity

Some of the other indicators in *A Nation at Risk* illustrate perhaps the most common misuse of statistics: selecting a statistic that, although accurate in itself, paints an incomplete and misleading picture. For instance, the report claimed that "there was a steady decline in science achievement scores of U.S. 17-year-olds as measured by national assessments in 1969, 1973, and 1977." This was true.

But the statement refers only to science, and only to 17-year-olds. What about the 9- and 13-year-olds also tested in national assessments? No

When comparing groups, it's important to make sure the groups are comparable.

rapid urbanization.) Using industrial output, Russia is first, China second, and the United States fifth. (Russian and Chinese industries are not as clean as U.S. industries.)

Which statistic is best? All of them together. Using only one would be like evaluating a center fielder only on his batting average or a quarterback only on yards gained per pass. You need more than one statistic to paint a complete picture.

Similarly, a recent e-mailer asked me whether a preschool program, which produced a four-month gain in vocabulary and math and cost \$6,000 per kid per year, was worth it. I said that I couldn't tell. For one thing, the program likely produced health, socialization, and other outcomes besides the two mentioned. In addition, the real value of the program might not be clear for years: It took long-term evaluations of the outcomes of the Perry Preschool Project and the Chicago Family Centers project to establish that society gained

remain remarkably trusting when it comes to statistics. Best (2001) observes that "Most of the time, most people simply accept statistics without question" (p. 4). This acceptance would be dangerous at any time, but given today's polarized politicization of education (and virtually everything else), it is particularly hazardous now. Educators can avoid this danger by following some basic principles of data interpretation.

Go Back to the Data

Many people call the National Commission on Excellence in Education's 1983 report *A Nation at Risk* "the paper Sputnik" because it focused attention on education in the same way Sputnik did in 1954. Some still refer to it today as a "landmark" study. It's a landmark, all right: a golden treasury of selected, spun, distorted, and even manufactured statistics.

After opening with a blast of Cold-Warrior rhetoric, the good commis-

“steady decline” in science for them. What about math? What about reading? No hint of any decline in either subject for any of the three age groups (National Center for Education Statistics [NCES], 2000).

The commissioners had nine trend lines available from NCES data (three ages times three subjects). Only one could be used to support crisis rhetoric, and that was the only one the commissioners mentioned.

Compare Rhetoric with the Numbers

Perhaps the most dangerous statistic is the one that Joel Best calls the *mutant* statistic. This statistic begins life as a legitimate datum, but mutates into something new and wrong. Best (2001) gives the example of the claim, widely circulated, that 150,000 women die in the United States each year from

Those who collect statistics have a purpose—sometimes benign, sometimes not.

anorexia. The U.S. Census Bureau's *Statistical Abstract of the United States* shows that 55,000 women ages 15-44 die each year of all causes. Even if anorexia had killed all 55,000, given that anorexia mostly affects young women, it is unlikely that we can find another 95,000 anorexia victims younger than 15 and older than 44. In fact, the proper statistic is that 150,000 women *suffer* from anorexia—and even this number is probably a bit inflated because it was produced by an activist group attempting to call attention to the problem.

Mutant statistics afflict education data as well. *Washington Post* pundit George Will wrote in one column that

FIGURE 1. Pass Rates and Average Scores Tell a Different Story

Pass Rate—Gap closed by 10 points			Average Scores—Gap increased by 8 points		
	2004	2005		2004	2005
Black students	60%	70%	Black students	62	68
White students	100%	100%	White students	78	92
Gap	40	30	Gap	16	24

These are hypothetical data. Score needed to pass = 60.

almost half of Chicago's public school teachers sent their own children to private schools (Will, 1993a). This was true. The figure was 43 percent at the time, and that was the highest proportion in the United States. (Religion figured strongly in the Chicago teachers'

Will's brain might have been addled by the work of Denis Doyle, whose reports using data from the 1980, 1990, and 2000 censuses have promoted the idea that public school teachers do send their kids to private schools in larger numbers than the general public does (Doyle, 1995; Doyle, DeSchryver, & Diepold, 2004; Doyle & Hartle, 1986). Doyle refers to teachers as “connoisseurs” of education, implying that if they send their kids to private schools, they must know something that the rest of us don't. He writes,

With teachers choosing private schools, the truth is self-evident: While they work in public schools, they choose private schools for their own children because they believe they are better. (1995)

This statement creates the impression that all public school teachers in all types of communities use private schools. But if we look beyond the rhetoric to the actual statistics, we find these figures for the United States as a whole (Doyle, 1995; Doyle, DeSchryver, & Diepold, 2004):

	Teachers	General Public
1990	12.1 percent	13.1 percent
2000	10.6 percent	12.1 percent

decisions.) But over a period of six months, Will's neurons replaced “Chicago” with “the nation”; in another column, he wrote, “Nationally about half of urban public school teachers with school-age children send their children to private schools” (Will, 1993b). This was not true. According to data from the 2000 Census, 17.5 percent of all urban families and 21.5 percent of urban public school teachers send their children to private schools. The rate ranges from 43.8 percent of teachers in the Philadelphia/Camden metro area down to 1.7 percent in Oklahoma City. In 21 of these top 50 cities, teachers use private schools less than urban families do (Doyle, DeSchryver, & Diepold, 2004).

The numbers show that teachers made less use of private schools than the general public did. What's more, despite all the lionization of private schools and the demonization of public schools during the 1990s, a smaller proportion of both teachers and the general public had children in private schools in 2000 than in 1990.

Make Sure That Groups Are Comparable

The statistics on the percentages of children sent to private schools point to another principle of data interpretation: When comparing groups, make sure the groups are comparable. Teachers and the general public are not comparable. Teachers are more likely to have at least a bachelor's degree and less likely to live below the poverty line. We need to consider the implications of these and similar factors before we draw conclusions about the two groups' public school-private school choices.

This principle often comes into play in figuring out the impact of high-stakes graduation tests. In 2004, Massachusetts announced that 96 percent of its seniors had passed the state test and would graduate. This was true, but it was true only for people who had begun the 2003-2004 school year as seniors and who were still in school. Many in the class of 2004 were no longer present and accounted for. When that cohort of students started 9th grade, it contained 78,000 students; by the time it reached 12th grade, there were only 60,000. Eighteen thousand students had decamped (Wheelock, 2004).

We don't know what happened to these students. Some, of course, left the state and might well have passed the test and graduated if they had remained. But others were retained in

StatusQuote

It is impossible for a man to learn what he thinks he already knows.

—Epictetus

grade and were no longer in the class of 2004. Some failed and dropped out or sought a General Equivalency Diploma. If we look at how many who started as 9th graders in the class of 2004 eventually graduated, we find rates ranging from 54 percent for Latino students to 80 percent for white students (Wheelock, 2004). We can't draw an accurate conclusion about the effects of high school graduation exams unless we consider all the groups, including those that did not graduate on time.

Know the Difference Between Rates and Scores

The Massachusetts example also illustrates another principle of data interpretation: Be aware of whether you are dealing with rates or scores. The two metrics can paint very different pictures of a situation. These days, most states are reporting some kind of rate: percent passing, percent proficient, or percent meeting state standards. But if we focus only on the proficiency cutoff, it doesn't matter whether the student exceeds it by one question or 40. We're looking at how many kids can jump over the barrier, not at how high they jump.

Moreover, using passing rates instead of scores can obscure the fact that the white-minority achievement gap may be increasing. Consider the theoretical data in Figure 1. If we look only at passing rates, black students have reduced the gap from 40 percent to 30 percent. But if we look at scores, the gap has actually increased from 16 points to 24 points.

This discrepancy might not be so important if the passing score actually meant something in terms of performance in the real world. But it doesn't. These passing scores are totally arbitrary. Some readers might recall that in my recent report on the condition of public education (Bracey, 2005), I awarded a Golden Apple to a student in Ohio because he refused to take the Ohio Proficiency Tests. It was not his act of defiance that garnered him a prize; it was the reasons he gave:

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. [The student was admitted to several colleges.] (p. 140)

Do the Arithmetic

Here's a final principle of data interpretation to examine on your own: Do the arithmetic. In 1995, an article in an education periodical (not *Educational Leadership*) stated that "Every year since 1950, the number of American children gunned down has doubled." Sit down with a calculator and a sheet of paper on which you write in one column the years from 1950 to 1994. Then assume that one child was "gunned down" in 1950 and let the figure double for each successive year. Have fun. ■

References

- Barringer, F. (2006, Jan. 26). United States ranks 28th on environment, a new study says. *The New York Times*, p. A3.
Berrueta-Clement, J. R., Schweinhart, L. J.,

For over a decade, we have successfully taught thousands of teachers how to create a **CONSISTENT** environment where all children can succeed.



www.ClassroomManagement.com
Home of the premier eLearning course with Harry and Rosemary Wong



Seton Hall University Offers a Fast Track to a Traditional Ed.D.

"The professors are extremely knowledgeable and talented. The class schedule is completely pre-arranged for the two year program. With advance planning, I have been able to travel to Seton Hall at very reasonable rates. I feel very fortunate to have been part of Seton Hall's Executive Ed.D. Program.

Duane Klusmann
Assistant Superintendent, SBISD
Houston, Texas
Graduated May 04
Appointed Superintendent Dec. 04

Our accelerated Ed.D. program allows you to complete your doctoral studies in just 10 weekends and two four-week summer sessions over a two-year period.

This rigorous program is designed exclusively for K-12 education administrators who need a flexible schedule that will not interfere with career responsibilities. Our cohort approach builds on shared experiences and the expertise of seasoned faculty and nationally recognized authorities in the educational field.

Financial aid loans cover the entire program cost, regardless of financial need.

Now accepting applications for the April 2007 cohort



400 South Orange Ave., South Orange, NJ 07079
www.shu.edu

College of Education and Human Services 1-800-313-9833

For more information e-mail: execedd@shu.edu or go to: education.shu.edu/execedd

Barnett, W. S., Epstein, A. S., & Weikart, D. P. (1984). *Changed lives: The effects of the Perry Preschool Program on youths through age 19*. Ypsilanti, MI: High/Scope Press.

Best, J. (2001). *Damned lies and statistics: Untangling numbers from the media, politicians, and activists*. Berkeley, CA: University of California Press.

Bracey, G. W. (2005). The fifteenth Bracey report on the condition of public education. *Phi Delta Kappan*, 87(2), 138-153.

Doyle, D. P. (1995). *Where connoisseurs send their children to school*. Washington, DC: Center for Education Reform.

Doyle, D. P., DeSchryver, D. A., & Diepold, B. (2004, Sept. 7). *Where do public school teachers send their kids to school?* Washington, DC: Thomas B. Fordham Institute. Available: www.edexcellence.net/foundation/publication/publication.cfm?id=333

Doyle, D. P., & Hartle, T. W. (1986). *Where public school teachers send their children to school: A preliminary analysis*. Washington, DC: American Enterprise Institute.

National Center for Education Statistics. (2000). *NAEP 1999 trends in academic progress*. Report No. NCES-2000-469. Washington, DC: Author.

National Commission on Excellence in Education. (1983). *A nation at risk*. Washington, DC: Author. Available: www.ed.gov/pubs/NatAtRisk

Reynolds, A. J. (2001, May 21-June 2). *Age 21 benefit cost analysis of the Chicago Child-Parent Center program*. Paper presented to the Society for Prevention Research, Madison, Wisconsin.

Wheelock, A. (2004, June 8). *Massachusetts Department of Education "progress report" inflates "pass rates" for the class of 2004* [Online]. Available: www.massparents.org/news/2004/passrate_2004.htm

Will, G. F. (1993a, March 7). When the state fails its citizens. *The Washington Post*, p. C7.

Will, G. F. (1993b, Aug. 26). Taking back education. *The Washington Post*, p. A27.

Gerald W. Bracey is an independent researcher and writer living in Alexandria, Virginia; gbracey1@verizon.net. His most recent book is *Reading Educational Research: How to Avoid Getting Statistically Snookered* (Heinemann, 2006).