

What's Wrong with Teaching to the Test?

Mr. Posner speculates on the evolution of a species that has been educated only to succeed on standardized tests.



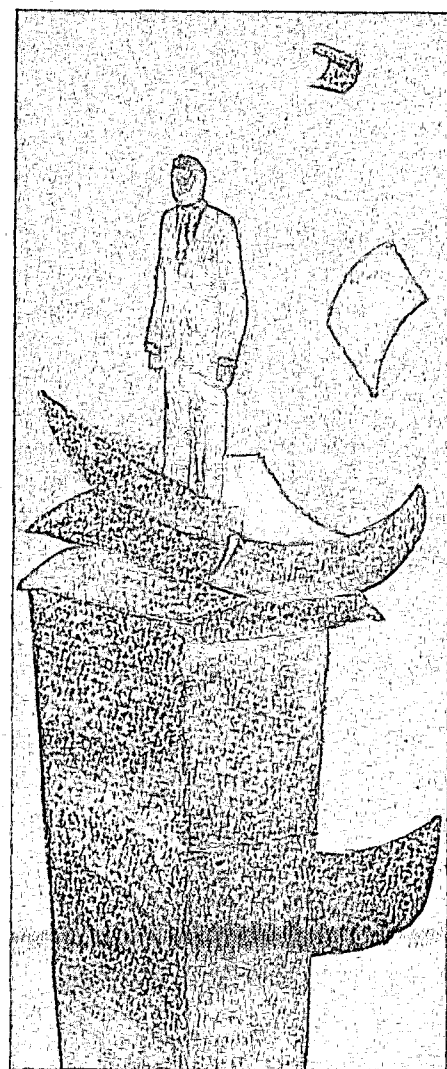
BY DAVE POSNER

IN CALIFORNIA, where I live, and in many other states, the quality of public education — and by extension the competence of its teachers — is being measured by students' scores on standardized achievement tests. The pressure on teachers and administrators to improve these scores is enormous. Up until the recent budgetary crises in California, teachers in districts whose test scores improved sufficiently (relative to national percentile rankings) were eligible for cash bonuses and extra money for programs. Though these "positive" incentives are gone, most of the "negative" incentives continue. Teachers in schools and districts whose scores fail to improve adequately are branded and subjected to various indignities. Schools that continue to fail to improve may be closed, and districts that continue to fail may be subject to state takeover.

Opponents of this so-called high-stakes testing complain that such intense pressure causes teachers to devote virtually all classroom time and resources to preparing students for the standardized test. This phenomenon is called "teaching to the test." Proponents of high-stakes testing respond that that is exactly as it should be. They argue that the tests measure success in teaching the curriculum and so "teaching to the test" is "teaching to the curriculum." And after all, isn't that what we want teachers to do?

I was led to consider this notion while thinking about the accomplishment of a former colleague who recently made a major breakthrough in a famous unsolved problem in mathematics (though he did not arrive at a complete solution). He has been working on this and related problems for more than 25 years, and some of these problems have been under attack for more than a century. I wondered whether the skills and mental processes necessary to attack problems of this magnitude were qualitatively different from those required to solve more routine problems or whether the intellectual requirements were essentially the same but applied over a much longer period.

DAVE POSNER, whose wife teaches third grade in a public school, is a former professor of mathematics and co-founder and chief technical officer of Encirq Corporation, San Francisco. Mr. Posner can be contacted at pdk@daveposner.com.



The kinds of problems that can appear on a standardized test are, of course, quite limited in form and complexity, as the student is allocated only a minute or two to complete each one. If the intellectual processes required to solve a really complicated problem are not essentially the same as those required to solve these simpler problems, then a student prepared only to solve

lems on standardized tests enable a student to attack this problem?

As stated, the problem might seem too vague. The student might well respond, "What do you mean by *should*?" But that is the way real problems usually confront us. Should the U.S. invade Iraq? Should I have sex? Should I smoke pot? Should I add this service or feature to my product line? Reducing these more or less vague problems to more concrete questions is a major part of the problem-solving process in the real world. We typically attack such "should" problems by analyzing the possible consequences of the proposed actions, along with the probabilities of those consequences and their relative costs and benefits. Does standardized test preparation enable

students to reduce a question about possible behaviors to a cost-benefit analysis?

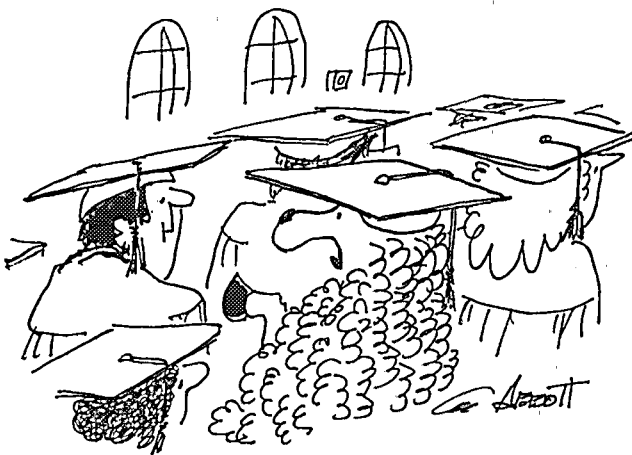
On a standardized test, all the data necessary to analyze a problem must be presented along with the problem. (Students are strongly discouraged from doing research during the test!) In contrast, for real problems, the necessary data for such an analysis are often either nonexistent, hidden, or questionable because they emanate from highly biased and conflicting sources. For the problem we are considering here, the last possibility probably applies. Education is a major economic and political activity, and you can be confident that many of the players, including test-makers, curriculum and textbook purveyors, staff development consultants, unions, politicians, the real estate industry, and even well-meaning outsiders (like me) have their own agendas or at least a bias of some sort. The research and design necessary to discover and evaluate such evidence are highly complex and error-prone and often require the use of analytical methods and an understanding of their appropriateness. What are reasonable measures of real-world problem-solving skills? Are there studies demonstrating the efficacy, relative to other activities, of standardized test preparation as a means of improving those skills? What are the data, and are they valid? There are certainly routine computational aspects to this process, but the really hard problem is to design, implement, and evaluate the process itself.

On a standardized test, the possible answers to a problem are limited and generally enumerated as a small multiple-choice list. For real problems, the list of possible outcomes is often enormous and at best partial. The discovery of these possibilities is essential to any meaningful analysis. In many cases the unanticipated

Individuals equipped only with the ability to solve routine problems would be those most vulnerable to displacement by automation.

standardized test problems could well lack the mental preparation required to attack really hard problems. Part of my concern about this matter is that routine problems are the most amenable to solution by computer. Thus individuals equipped only with the ability to solve routine problems would be those most vulnerable to displacement by automation.

Of course, solving famous unsolved problems in mathematics is a special calling and probably not a reasonable model for what we should expect from most of our students. As a model for evaluating whether teachers should teach to the test, we should use something more typical of the kind of everyday problems that concern us as workers or parents or citizens. But we needn't look far. The very question we are considering — Should teachers teach to the test? — strikes me as a typical example. Would the capabilities required to solve prob-



"Would you believe it? All this started one day when I followed Mary to school."

ed and unintended consequences of an action are the ones that matter most. Could that be the case for the present question? For example, could the creation of a system intended to improve educational quality result in a population unable to think beyond the superficial? Could the unintended result be a population so intellectually incompetent that it can't recognize its own incompetence?

For real problems, the appropriate methods of attack are not immediately obvious and may well vary greatly from those that apply to problems that seem similar. In contrast, on a standardized test, where there is no time for subtlety or deep analysis, problems are by necessity formulaic. Could an education driven by standardized test scores leave students unable to understand such subtleties? For example, students in a classroom constitute a set of problems for a teacher that are superficially similar but at a deeper level radically different. Could the inability to appreciate these distinctions be another unintended consequence of teaching based on standardized tests? Consider that the pressure on teachers to "teach to the test" is more accurately described as pressure to "teach to the standardized test metrics" by which the teacher's performance is measured. Optimizing such metrics requires ignoring individual students in favor of statistical abstractions. Such an approach leads naturally to the kinds of "operations research" methods used in business, in which resources are allocated not according to the needs of individuals but according to the needs of the abstractions. Such statistical optimization leads naturally to regimentation. McDonald's ensures the quality of its products and services by precise regimentation of its processes.

We now see this happening in education in the form of "scripted programs." In these programs, teaching behavior is regimented down to the exact material, timing, and wording of the instruction. Could our obsession with standardized tests reduce teaching itself to a simplistic and ultimately ineffective activity that would be amenable to automation?

I see the obsession with standardized tests in Darwinian terms. We are in effect putting our kids (and their teachers!) on an isolated atoll under the evolutionary force of a strange selection process based on standardized tests. The inevitable product of this process is a species that is as custom-engineered as any carbon-based life form can be to solve trivial problems. Like most exotic species, this one is unlikely to be able to adapt to and compete in the larger world. The irony is

that it is unlikely to prevail even in its chosen niche, where the fittest survivors will most likely be made of silicon. BK

Test Anxiety

(Continued from page 744)

treat their families, peers, and especially strangers with love, compassion, and forgiveness. The way we learn to live our lives as adults, and not our test scores, is the true measure of our schooling.

1. Adapted from Myra Sadker and David Sadker, "Standards Back Then," in idem, *Teachers, Schools, and Society*, 6th ed. (New York: McGraw-Hill, 2003), pp. 268-69.
2. Jay Heubert, "First, Do No Harm," *Educational Leadership*, December 2002/January 2003, pp. 26-30; Linda M. McNeil, "Creating New Inequalities: Contradictions of School Reform," *Phi Delta Kappan*, June 2000, pp. 728-34; and John Merrow, "Undermining Standards," *Phi Delta Kappan*, May 2001, pp. 653-59.
3. *Quality Counts 2001: A Better Balance* (Bethesda, Md.: Editorial Projects in Education, 11 January 2001).
4. James Hoffman, Lori Czop, and Scott Paris, "High-Stakes Testing in Reading: Today in Texas, Tomorrow?," *Reading Teacher*, vol. 54, 2001, pp. 482-92.
5. *Quality Counts 2001*.
6. Hoffman, Czop, and Paris, op. cit.
7. Audrey Amrein and David Berliner, "High-Stakes Testing, Uncertainty, and Student Learning," *Education Policy Analysis Archives*, March 2002, <http://epaa.asu.edu/epaa/v10n18/>.
8. Monty Neill, "The Dangers of Testing," *Educational Leadership*, February 2003, pp. 43-46.
9. Alfie Kohn, *The Case Against Standardized Testing: Raising Scores and Ruining Schools* (Portsmouth, N.H.: Heinemann, 2000).
10. Amrein and Berliner, op. cit.
11. Heubert, op. cit.; McNeil, op. cit.; and Neill, op. cit.
12. Norman Draper, "Testing Firm to Pay Students \$7 Million for Grade Goof," *Star Tribune Online*, 26 November 2002.
13. Diana B. Henriques, "Rising Demands for Testing Push Limits of Its Accuracy," *New York Times*, 2 September 2003.
14. Neill, op. cit.
15. Heubert, op. cit.
16. Amrein and Berliner, op. cit.; Heubert, op. cit.; and Neill, op. cit.
17. Derrick Jackson, "At Best, Silly, At Worst, Racist," in Kathy Swopes and Barbara Miner, eds., *Failing Our Kids: Why the Testing Craze Won't Fix Our Schools* (Milwaukee: Rethinking Schools, 2000), pp. 18-19.
18. N. R. Kleinfeld, "The Elderly Man and the Sea? Test Sanitizes Literary Texts," *New York Times*, 2 June 2002.
19. David Hoff, "Standards at Crossroads After Decade," *Education Week*, 22 September 1999, pp. 1, 9.
20. *Quality Counts 2001*.
21. John Goodlad, "Kudzu, Rabbits, and School Reform," *Phi Delta Kappan*, September 2002, pp. 16-23.
22. Tirupalavanam Ganesh, "Educators' Images of High-Stakes Testing: An Exploratory Analysis of the Value of Visual Methods," paper presented at the annual meeting of the American Educational Research Association, New Orleans, 2002. To see the teachers' drawings, visit Ganesh's website at <http://ganesh.ed.asu.edu/aims>. BK

A vertical bar on the left side of the page, consisting of a series of horizontal segments in shades of yellow and orange, with a small red diamond at the top.

COPYRIGHT INFORMATION

TITLE: What's Wrong with Teaching to the Test?
SOURCE: Phi Delta Kappan 85 no10 Je 2004
WN: 0415701226010

The magazine publisher is the copyright holder of this article and it is reproduced with permission. Further reproduction of this article in violation of the copyright is prohibited.

Copyright 1982-2004 The H.W. Wilson Company. All rights reserved.