

ISSUE BRIEF

www.centerforcsri.org

July | 2006

'Teach to the Test'? Just Say No

BY CRAIG D. JERALD

Every spring, education-related newspaper and magazine stories raise the alarm that schools are "teaching to the test." Scores of articles and editorials paint a disheartening picture of frustrated teachers forced to abandon good instructional practices for a relentless stream of worksheets based on boring, repetitive test-preparation materials. Even Hollywood actors are chiming in. Actress Alfre Woodard recently told a Louisiana newspaper, "My sister-in-law is left standing in front of her class with a pamphlet, teaching to the test because everyone must pass."¹

Although the phrase—and the concern—are hardly new, many observers blame the No Child Left Behind (NCLB) Act for escalating teaching to the test from a problem into an epidemic. The law "virtually transformed the concept of education," according to a recent editorial in the *San Francisco Chronicle*, "turning teaching and learning into a mere exercise in prepping students to test well."²

According to that logic, teaching to the test is as unavoidable as a force of nature, as inevitable as gravity. And the choice between good instructional practice and good test scores is really no choice at all, since those who opt not to bow to the pressure will reap harsh consequences under tough accountability systems.

Such claims often are taken at face value. But what do we really know about the phenomenon? Does high-stakes testing always force educators to "dumb down" instruction to focus on rote skills and memorization? Do schools that spend a lot of time on test preparation and "drill and kill" instruction actually perform better on standardized tests than those that do not? Those might sound like easy-to-answer questions, but the answers to those questions are surprising. Many forms of teaching to the test are as unnecessary as they are harmful.

What's Wrong With Teaching to the Test?

The phrase "teaching to the test" is used widely but seldom defined, causing much confusion about what it means and whether it is bad or good. Indeed, in a

recent editorial in the *Washington Post*, the respected education reporter Jay Mathews claimed that teaching to the test simply means aligning classroom instruction and curriculum to standards; the practice is a good one that should be supported.³ Teachers rushed to tell Mathews that what he described was not the kind of teaching to the test that they and their colleagues are worried about.

Assessment expert W. James Popham helps to clarify the difference. He defines two kinds of assessment-aware instruction: “curriculum-teaching” and “item-teaching.”⁴ *Curriculum teachers* focus on the full body of knowledge and skills represented by test questions even though tests can employ only a sample of questions to assess students’ knowledge about a topic. For example, if students will be tested on fractions, curriculum teachers will cover range of knowledge and skills related to fractions so students understand what fractions are, know how to manipulate them mathematically, understand how to use them to solve more complex problems, and are able to communicate with and about them. *Item teachers* narrow their instruction, organizing their teaching around clones of the particular questions most likely to be found on the test—and thus teach only the bits of knowledge students are most likely to encounter on exams. For example, item teachers might drill students on a small set of vocabulary words expected to be assessed rather than employing instructional strategies that help students develop the kind of rich and broad vocabulary that best contributes to strong reading comprehension.

The latter kind of teaching to the test is unethical. For one reason, it misrepresents how much students really have learned about a topic. In the example, students who learned only the 10 words on the vocabulary portion of the reading test will score well even though they have not developed a broad vocabulary, which is supposed to be the goal. In mathematics, students who have been

drilled only on testlike questions do not have the opportunity to master a particular skill or concept and often cannot correctly answer questions that assess the same skill or concept in a different way. For example, one study found that in a district that relied heavily on item drilling, 83 percent of students selected the correct answer to a multiple choice item written as “ $87 - 24 =$.” However, only 66 percent could provide the correct answer to the open ended item “Subtract 24 from 87.”⁵

Popham also contends that “because teaching either to test items or to clones of those items eviscerates the validity of [tests]... item-teaching is reprehensible. It should be stopped.”⁶ But the problems with teaching to the test go beyond the fact that it interferes with test validity. Parents and educators are much more concerned with how it affects the curriculum and classroom instruction itself.

For example, some worry that item teaching and other test-preparation strategies are taking over more of the weeks and months prior to testing. “They are losing a week of instruction to testing, which is bad enough,” lamented a commentator in the *Chicago Sun-Times* last March. “But the test week comes on top of two or more weeks spent teaching kids how to take the test effectively.”⁷ Others worry that the negative effect on instruction stretches back to August and September, with “drill and kill” strategies that substitute memorization for understanding and strangle good instruction all year long.

According to Lauren Resnick and Chris Zurawsky, the combination of accountability, the lack of a clear curriculum, and cheaper off-the-shelf tests is a recipe for bad teaching. “When teachers match their teaching to what they expect to appear on state tests of this sort,” they write, “students are likely to experience far more facts and routines than conceptual understanding and problem-solving in their curriculum.... Narrow tests...can become the de facto curriculum.”⁸ Resnick and

Zurawsky do not object to accountability per se, but warn that it can lead to inappropriate instruction if it is not backed up with strong curricula and aligned assessments.

Like Resnick and Zurawsky, many observers worry that drill-focused forms of teaching to the test can crowd out opportunities to teach students more advanced cognitive skills, such as how to solve problems and communicate effectively. They point to the work of economists, such as Frank Levy and Richard Murnane, who warn that all kinds of jobs, but particularly higher paying jobs, increasingly require fewer rote and routine skills and ever more complex skills. Analyzing tasks performed in jobs across the economy between 1969 and 1999, for example, Levy and Murnane found a big decline in rote tasks and routine work along with a skyrocketing demand for “expert thinking” skills (the ability to solve problems that require more than simply following rules or applying knowledge to new situations) and “complex communication” skills.⁹

Levy, Murnane, and other economists argue that young people who are denied the opportunity to develop such advanced skills will be at an increasing disadvantage in the changing economy of the 21st century.¹⁰ That means educators who settle for “drill and kill” instruction—or who do not at least balance such instruction with more complex assignments—will be trading long-term benefits to students for short-term gains on standardized tests.

Fool’s Gold and False Choices

The decision to narrowly teach to the test might be bad for students in the long run, but is it really inevitable? Is there an unavoidable trade-off between helping students develop advanced problem-solving and communication skills they will need later in life and helping them perform better on standardized tests

while they are in school? More to the point, do “drill and kill” strategies for teaching to the test actually produce higher test scores than other forms of instruction?

During the 1990s, Chicago instituted a number of accountability policies requiring students and schools to meet performance standards on nationally norm-referenced standardized assessments. Some educators and reformers worried that the policies would force teachers to abandon assignments that ask students to engage in critical thinking in favor of strategies focused on memorization, drill, and repetitive practice. A trio of widely respected researchers—Newmann, Bryk, and Nagaoka—then affiliated with the Chicago Consortium on School Research, decided to investigate that assumption, asking “What happens to students’ scores on standardized tests of basic skills when urban teachers...assign work that demands complex thinking and elaborated communication...?”¹¹

The researchers conducted a three-year study analyzing classroom assignments and student gains on standardized tests across more than 400 Chicago classrooms in almost 20 elementary schools. Nearly 2,000 classroom assignments were scored based on a rubric that evaluated the extent to which the assignments called for “authentic intellectual work” from students—applying basic skills and knowledge to solve new problems; expressing ideas and solutions using elaborated communication; and producing work related to the real world beyond the classroom. Note that the definition of authentic instruction does not simply mean “creative” assignments that ask students to use their imaginations but rather “disciplined inquiry,” in which students apply imagination and logic as *well as* “the basics”—vocabulary words, facts, algorithms—to complete tasks that go beyond answering multiple-choice questions.

For example, one sixth-grade mathematics assignment that scored high on the



researchers' rubric asked students to assume they had \$10,000 to invest in a stock. Students selected and tracked their own stocks, reading the newspaper and using their knowledge of fractions to calculate gains and losses. At the end of 10 weeks, students decided whether to buy more of the stock or to sell it, and they presented an oral report describing the results of their investment and their decision to unload or reinvest. In contrast, a low-scoring assignment asked sixth graders to complete a worksheet adding or subtracting pairs of simple fractions, such as $4/5 - 2/5$.

The researchers called the more advanced assignments "authentic" precisely because they were thought to more closely mimic the kinds of tasks adults perform in their jobs. And, indeed, although the rubrics were developed well before Levy and Murnane conducted the economic study described earlier, their criteria for evaluating classroom assignments seem to closely parallel the kinds of "expert thinking" and "complex communication" skills the two economists found to be in ever greater demand in today's workplace.

Newmann, Bryk, and Nagaoka then analyzed student test-score gains on the commercially developed, nationally norm-referenced Iowa Test of Basic Skills (ITBS) assessment and the state-developed Illinois Goal Assessment Program (IGAP) exams. The results were startling. In classrooms where teachers employed more authentic intellectual instruction, students logged test-score gains on the ITBS that exceeded the national average by 20 percent. However, students who were given few authentic assignments gained much *less* than the national average. A similar pattern emerged when researchers examined results on the IGAP assessments.

To be sure of their findings, the researchers took into account cross-classroom differences in students' prior-year test scores as well as race, gender, and poverty levels. They also

conducted a check to examine whether more advanced students disproportionately were getting the more demanding assignments. Surprisingly, the answer was no—the distribution of highly authentic assignments was determined by teachers' own "dispositions and individual choices" rather than by the designated level of students they taught. Moreover, both high- and low-achieving students benefited from the more demanding, authentic assignments.

Those results strongly suggest that accountability and standardized tests *need not be* in conflict with good instruction, and that Resnick and others are wrong to assume that off-the-shelf tests *require* teachers to give up teaching higher level skills. "Fears that students will score lower on conventional tests due to teacher demands for more authentic intellectual work appear unwarranted," the researchers concluded. "To the contrary, the evidence indicates that assignments calling for more authentic intellectual work actually improve student scores on conventional [standardized] tests."¹² In other words, teaching to the test by "dumbing down" instruction offers only a kind of fool's gold, promising a payoff that it does not deliver. The choice between good instruction and good test scores is a false one.

The researchers hypothesized that using basic skills to perform complex intellectual tasks actually helps students better internalize such skills and apply them across a wide range of tasks, including standardized tests. However, they also cautioned that no one instructional strategy can serve all purposes. For example, some students might need to practice basic fraction problems before moving on to the more complex project involving stocks. Thoughtful teachers employ a variety of strategies to ensure that students develop basic skills and can apply those skills to complex tasks grounded in real-world challenges.

Many experts also agree that some forms of direct test preparation can be healthy in small doses, and it might even be necessary for tests to provide valid results. For example, students unfamiliar with the test-question format might need help understanding how to answer certain kinds of items so they truly can show what they know. However, a little teaching about test format goes a long way, and engaging in more test preparation than absolutely necessary can depress scores, since it takes time away from the kinds of classroom assignments that help students master the content the test will assess.

Making Better Choices

Clearly, it is possible for educators to make better choices about how and when to teach to the test than the alarmist newspaper articles and editorials would seem to suggest. Cynthia Kuhlman, principal of Atlanta's high-poverty, high-minority Centennial Place Elementary School, where nearly all students consistently meet state standards on Georgia's assessments, says, "We don't teach to the test here at all. We have a curriculum that is mapped to the state's standards, and we teach almost entirely through theme-based projects. You would be hard pressed to find a worksheet at Centennial Place."¹³

At University Park Campus School, a high-poverty secondary school in Worcester,

Massachusetts, not only did all 10th graders pass the state exit exam last year, four out of five scored at the *advanced* or *proficient* level on it, compared with only about a third of students statewide. Yet the school has no prepackaged instructional reading program, and teachers push students to think critically about what they read, using quotes from the text to support their opinions. "So, despite the evidence of their testing success, University Park is not training test-takers," a columnist for *The Providence Journal* recently concluded. One teacher told her that reason was simple stubbornness.¹⁴

But some schools will need more than simple stubbornness to resist the lure of teaching to the test. Many teachers and administrators clearly do feel pressure to engage in "item-teaching" and rote instruction; and, especially in states that use off-the-shelf norm-referenced exams, educators increasingly worry that they might be sacrificing higher scores if they do not. It is time to overturn the common assumption that teaching to the test is the *only* option schools have when faced with high-stakes testing. Overreliance on "drill and kill" and test-preparation materials is not only unethical in the long-term but ineffective in the short-term. Because there really is no trade-off between good instruction and good test scores, this is that rare case when educators can have their cake and eat it, too.

This is the third in a series of issue briefs to be written for The Center for Comprehensive School Reform and Improvement during 2006. These commentaries are meant to help readers think beyond simple compliance with federal law or basic implementation of programs: What unacknowledged challenges must educators and leaders confront to help schools operate more effectively and to sustain improvement over the long run? In what ways does the conventional wisdom about teaching, learning, and school improvement run counter to current research and get in the way of making good decisions? What are the emerging next-generation issues that educators will face next year and five years from now? Readers can visit www.centerforcsri.org to obtain other papers in this series and to access additional information on school reform and improvement.

Endnotes

- ¹ Kleinschrodt, M. H. (2006, April 8). Lessons learned. *The Times-Picayune*. Retrieved June 30, 2006, from <http://www.nola.com/printer/printer.ssf?/base/entertainment-0/1144478417309480.xml>
- ² O'Neill, J. (2006, March 12). Leaving creativity behind: Drilling for tests kills curiosity and imagination. *San Francisco Chronicle*, p. E-5. Retrieved June 30, 2006, from <http://www.sfgate.com/cgi-bin/article.cgi?file=/chronicle/archive/2006/03/12/ING2DHLDO01.DTL>
- ³ Mathews, J. (2006, February 20). Let's teach to the test. *The Washington Post*, p. A21. Retrieved June 30, 2006, from <http://www.washingtonpost.com/wp-dyn/content/article/2006/02/19/AR2006021900976.html>
- ⁴ Popham, W. J. (2001, March). Teaching to the test? *Educational Leadership*, 58(6), 16–20.
- ⁵ Shepard, L. A. (1997). *Measuring achievement: What does it mean to test for robust understanding?* Princeton, NJ: ETS.
- ⁶ Popham, W. J. (2001, March). Teaching to the test? *Educational Leadership*, 58(6), 16–20.
- ⁷ Richards, C. (2006, March 15). 'No Child' law is flunking the test. *Chicago Sun-Times*. Retrieved June 30, 2006, from <http://www.suntimes.com/output/richards/cst-edt-cindy151.html>
- ⁸ Resnick, L., & Zurawsky, C. (2005, spring). Getting back on course: Standards-based reform and accountability. *American Educator*. Retrieved June 30, 2006, from http://www.aft.org/pubs-reports/american_educator/issues/spring05/resnick.htm
- ⁹ Levy, F., & Murnane, R. J. (2004). *The new division of labor: How computers are creating the next job market*. Princeton, NJ: Princeton University Press.
- ¹⁰ Olson, L. (2006, March 22). Economic trends fuel push to retool schooling. *Education Week*, 25(28), 1, 20, 22, 24. Retrieved June 30, 2006, from <http://www.edweek.org/ew/articles/2006/03/22/28prepare.h25.html>
- ¹¹ Newmann, F. M., Bryk, A. S., & Nagaoka, J. K. (2001). *Authentic intellectual work and standardized tests: Conflict or coexistence?* (p. 10). Chicago: Consortium on Chicago School Research. Retrieved June 30, 2006, from <http://www.consortium-chicago.org/publications/pdfs/p0a02.pdf>
- ¹² Newmann, F. M., Bryk, A. S., & Nagaoka, J. K. (2001). *Authentic intellectual work and standardized tests: Conflict or coexistence?* (p. 29). Chicago: Consortium on Chicago School Research. Retrieved June 30, 2006, from <http://www.consortium-chicago.org/publications/pdfs/p0a02.pdf>
- ¹³ Mathews, J., & Chenoweth, K. (2006, April 4). Much better than adequate progress. *The Washington Post*. Retrieved June 30, 2006, from <http://www.washingtonpost.com/wp-dyn/content/article/2006/04/04/AR2006040400644.html>
- ¹⁴ Steiny, J. (2006, March 19). Failure wasn't an option (Edwatch). *The Providence Journal*. Retrieved June 30, 2006, from http://www.projo.com/education/content/projo_20060319_edwatch19.3357a24.html

Administered by Learning Point Associates in partnership with the Southwest Educational Development Laboratory (SEDL) and WestEd, under contract with the Office of Elementary and Secondary Education of the U.S. Department of Education.



P: 877-277-2744 < W: www.centerforsrl.org

1100 17th Street NW, Suite 500
Washington, DC 20036