

The Teacher, Standardized Testing, and Prospects of Revolution
Author(s): Robert E. Stake
Source: *The Phi Delta Kappan*, Vol. 73, No. 3 (Nov., 1991), pp. 243-247
Published by: Phi Delta Kappa International
Stable URL: http://www.jstor.org/stable/20404603
Accessed: 21/03/2013 17:16

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Phi Delta Kappa International is collaborating with JSTOR to digitize, preserve and extend access to *The Phi Delta Kappan.*

http://www.jstor.org

This content downloaded from 159.242.252.8 on Thu, 21 Mar 2013 17:16:20 PM
All use subject to JSTOR Terms and Conditions

The Teacher, Standardized Testing, and Prospects Of Revolution



Illustration by Mario Noche

If we are to counterbalance the simplification of education brought about by testing, we will have to rely on the conceptual powers of classroom teachers, Mr. Stake observes.

By Robert E. Stake

EVEN A QUICK glance at the President's recent call for an educational revolution reveals the lack of a central place in the strategy for the classroom teacher. The proposal mentions evaluating teachers' skills and paying them according to merit but says nothing about relying on the ingenuity and experience of American teachers. Perhaps this was a political oversight. It was definitely a design flaw. If we are to counterbalance the simplification of education brought about by testing (an effect documented by other articles in this section), we will have to rely more on the conceptual powers of classroom teachers.

ROBERT E. STAKE is director of the Center for Instructional Research and Curriculum Evaluation at the University of Illinois, Urbana-Champaign. This article is based in part on a monograph titled Validity and Invalidity of Standardized Mathematics Achievement Testing, *prepared by the National Center for Research in Mathematical Sciences Education at the University of Wisconsin, Madison.*

This content downloaded from 159.242.252.8 on Thu, 21 Mar 2013 17:16:20 PM
All use subject to JSTOR Terms and Conditions

TEACHER-MADE ASSESSMENTS

In the 1990s education continues to be labor-intensive. Attempts over the years to automate teaching have been largely unsuccessful.[1] Why do we continue to put at least one expensive laborer in each classroom of every school? It is not because teacher unions are featherbedding.

Classroom teaching varies from room to room because diversity is the norm.

It is not because it takes a scholar to maintain discipline. It is not mainly to ensure the presentation of specific subject matter. We place almost three million teachers in American classrooms because *managing* the conditions for learning requires almost constant attention.[2]. It requires, for example, a constant watch to recognize students' readiness to learn, to identify their individual characteristics, and to notice obstacles and intrusions that interfere with their learning – in short, a continual assessment of student achievement.[3]

In the ordinary practices of schooling, teachers' assessment summaries are not used to communicate the scholastic integrity of the school or the career prospects of students. Such reporting is beyond the present and probably the future skill of teachers. As indicated in studies of teaching by Dan Lortie and others, teachers' informal assessments are used to direct the activities of learners and to reallocate time-on-task.[4] This kind of assessment has little to do with a science of education or with formal testing; rather, it is an intuitive art, which is developed through day-to-day teaching and which sometimes matures into disciplined reflection.

COMPLEXITY OF EDUCATION

During the past 40 years, various crusades to reform education have fallen short. One of the problems has been that practicing teachers and reformers hold such different notions of teaching and learning. For each subject and for each class or grade, teachers deal with enormous inventories of facts, relationships, exercises, and skills – mostly the distillation of previous teaching.[5] Of course, most of these components of subject matter can be classified according to the broadsides of reformist literature, but guidelines for reform are typically weak in their specifications of what and how to teach.

It is not difficult to set state or national goals or standards – and it is important to do so. But even massive collections of such statements give a terribly incomplete definition of education, enough so that a teacher striving vigorously to satisfy any stated goal is drawn away from the internal logic of the course and jeopardizes its coherence within student experience.[6] Each goal or standard should be regarded as an umbrella under which many teachings are appropriate and many learnings relevant. Test-item pools, curriculum guides, and textbook chapters are additional umbrellas. What actually occurs under the umbrellas, each teacher decides – often on the spot, often differently for different students. The essences of instruction – e.g., emphasizing a particular skill, covering a good number of topics while allowing for an appropriate amount of redundancy, drawing on the uniqueness of previous experience, informally assessing student progress – are but vaguely guided by statements of goals and standards.

Many governors, newspaper writers, and educators believe that more of teaching should be determined in advance and standardized across classrooms.[7] To an extent, the major goals of any school are prespecified and common. But education remains a highly individual experience. Classroom teaching varies from room to room for a good reason: diversity is the norm. Schools are different, teachers are different, children are different. We educators and researchers are not smart enough yet to write prescriptions for het-



"I kept saying, 'Maybe we should clean off the front of the refrigerator' – but, nooo . . ."

This content downloaded from 159.242.252.8 on Thu, 21 Mar 2013 17:16:20 PM
All use subject to JSTOR Terms and Conditions

erogeneous schools and children. Our formal plans are embarrassingly simplistic when compared to the routine and intuitive conceptualizations held by teachers. To organize each course of study to fit statistical standards would require more than a revolution – it would mean overthrowing all serious notions of education. Teaching has developed as an art; as a technology it is far less sophisticated.

When we analyze what a teacher is doing, we find topics and activities connected in logical ways. The topics of trigonometry are closer to the topics of geometry than to the topics of arithmetic. The idea of percentage is more closely related to fractions than to probability. Indirectly more often than directly, the teacher amplifies the organization of the textbook and builds the epistemological relationships. The math teacher is not working so much to develop a general mathematical ability in youngsters as to develop their knowledge of specific topics and their skills at solving specific kinds of problems.

When teaching language arts, the teacher is not aiming to make children literate or (unless harassed) to have them score higher on a literacy test, but rather to help them gain command of the components of language and communication (e.g., modifiers, the language of comparatives and superlatives, metaphors, cognates, British versus American spellings, Shakespearean style, and thousands of other specifics gathered under the literacy umbrella). The teacher sees education in terms of mastery of specific knowledge and sophistication in the performance of specific tasks, not in terms of literacy or the many psychological traits commonly defined by our tests.

The technology of *representing* pedagogy and epistemology is not highly developed. Classification systems and content/skill grids are common in curriculum offices, but there are few devices to portray conceptual links between topics and the appropriateness of moving pedagogically from one kind of content to another. Yet, just as ancient travelers reached destinations before there were maps, teachers teach without maps, without blueprints. Intuitively, good teachers merge topical streams, capitalize on personal experience, and draw out and preserve a youngster's line of thought.

But shouldn't we have blueprints *and* intuition? Yes, but they compete, and one or the other needs to be in charge. Teaching does not move easily back and forth between the pursuit of understanding and the pursuit of test performance. We are forced to choose between a strategy that optimizes teacher conceptualization and a strategy that maximizes attention to national goals.

Teachers have long been dubious about claims for elevating student achievement via centrally mandated tests.

REFORM BASED ON TEST INFORMATION

In America 2000 Secretary of Education Lamar Alexander indicated that, to have an effective education system, we must know how much each child knows. He suggested that parents have a right to know whether or not their child understands what is needed to be a competitive worker for the world marketplace or what is needed to be a scientist in the 21st century.

We should not promise to deliver what we are not close to knowing. Our tests do not tell us what students know; they tell us which students know the most about the particular questions asked and which students will do the best on future scholastic assignments. Our tests provide valid generalizations about how students stack up against one another. Information about the quality of education is *not* what our tests provide. A new national test will not provide that information. It probably will not help teachers teach – particularly if it continues to be the case that the cost of administering anything other than paper-and-pencil, machine-scorable, group tests largely precludes the use of a more "authentic assessment" at the state or national level.[8]

Over the decades, research studies – most recently one conducted by Carol Tittle, Kathy Kelly-Benjamin, and Joanne Sacks of the City University of New York – have made it clear that teachers do not find standardized achievement test scores very useful.[9] The tests seldom identify student talents that teachers had not recognized previously. They seldom provide teachers with diagnostic information that helps redirect their teaching. They are of little help as pretests to gauge the prior understanding of students on a topic about to be introduced.[10]

However, teachers recognize a distinction between testing as information gathering and testing as instructional management. Teachers do find standardized testing to be useful as a process for orienting students and teachers to a common curriculum. In a national survey of my own, a majority of math teachers judged there to be more good than harm in standardized testing. While they acknowledged the harmful pressures testing can exert and said they use the information very little, they noted that they capitalize on the anticipation. The upcoming test frightens the conscientious student, assuages the concerned parent, protects the authority of the teacher. Does testing change what teachers teach? Most said, "A little." Only about one teacher in eight spoke of a serious deterioration in teaching because of gearing content and class time toward tests.[11]

Teachers have long been dubious about claims for elevating student achievement via centrally mandated tests. In 1981 the state of Florida had perhaps the most aggressive state testing program in the nation. State Superintendent Ralph Turlington repeatedly indicated that Florida teachers were solidly behind the testing program. At the time, I was working on a national evaluation team that happened to be studying sex equity in the nation's 10th-largest school district, the Broward County (Florida) Schools. We asked a 15% sample of teachers: "In this district's schools, how much are the following interfering with students' getting a good education? Racial discrimination; discrimination according to sex; bilingual problems; overemphasis on testing." About half of the teachers indicated that testing interfered more than the other

This content downloaded from 159.242.252.8 on Thu, 21 Mar 2013 17:16:20 PM
All use subject to JSTOR Terms and Conditions

three factors.[12] This is but one example of the attitude that has long prevailed among teachers: they have *essentially no confidence in testing as the basis of the reform of schooling in America.*

A SCIENCE OF EDUCATION

We cannot use standardized achievement tests to make good on Secretary Alexander's promises unless we can connect performance on the tests to a science of education. We cannot reform schooling through testing unless we know what to do when students get low scores. Should we redirect teaching time to focus on the missed items at the risk of further diminishing students' range of experience, complexity of thinking, engagement in interesting problems? Our research does not say. A science of education is only barely begun. The most systematic knowledge of teaching and learning exists in the minds of teachers, augmented in a number of ways by the findings of research and the schemes of technicians. This combination is not yet something to be called a science of education, not something to justify the secretary's claim.[13]

We should be wary of premature calls for the use of technology to manage instruction, for management by objectives, for management by statistical indicators.[14] Look to Vietnam. The deceit of waging war by the strategic use of statistical indicators was illustrated by Neil Sheehan in his biography of Col. John Vann.[15] Look to Detroit. According to David Halberstam, the loss of American dominance in the automobile market followed the ascendance to executive command of economists and bankers.[16] Some observers of the workplace argue that worker empowerment – giving workers a "real" conceptual role – increases productivity and corporate health.[17] Secretary Alexander should make the same case for teachers.

UNDERMINING EDUCATION

Our field studies of American classrooms inform us that the American teacher remains a major asset – perhaps not as capable as we would like or all that children deserve, but largely pleasing to the local community and to school authorities – more the artist and even more the technician than reformist agitation would suggest. Most teachers have heard the call to reform, are sympathetic to it, and are hopeful of contributing to improved student assessment. Many are troubled by the diversion of instructional time to preparation for testing. Most do not see that mandated assessment is already changing the nature of education in America.

Research shows that it is a mistake to design a revolution in American schools around a national testing program.

Education is being redefined. Standardized testing – intentionally, with noticeable and often harmful effects – does change education.[18] Teachers report that, with increased testing and standardization of curriculum, they attend more to the so-called basics, the most elementary knowledge and skills, and less to the deep understanding of even a few topics. The dangers in current school reform are several: overstandardization, oversimplification, overreliance on statistics, student boredom, increased dropouts, a sacrifice of personal understanding, and, probably, a diminution of the diversity of intellect among people.

Specifying standard academic skills and curricular topics for all to master has not upgraded education. There are other roads to reform – some that give children increased opportunities to confront intellectual problems, to voice perplexity, and to propose explanations.[19] Many of us see it as essential that individual children be helped to relate their studies to personal experience. To those who hold this view, state reforms, in trying to raise standards, have relied overmuch on common goals and common test performance. We could not do without common aspirations and expectations, but there is also a great need for individualized teaching and for eliciting personal interpretations from each child. Many of our teachers have the necessary abilities. Overemphasis on common goals diverts their efforts. Andrew Porter discusses one possibility for reconciling the need for common standards with the need for individualization:

> Simply telling teachers what to do is not likely to have the desired results. Neither is leaving teachers alone to pursue their own predilections. But . . . it might be possible to shift external standard setting away from reliance on rewards and sanctions . . . and toward reliance on authority. . . . One approach to building authoritative standards would be to involve teachers seriously in the business of setting standards. . . . Through the process of teacher participation . . . the standards would take on authority.[20]

I have not intended to glorify the American teacher. If you believe I am saying that professors, politicians, pundits, and parents should shut up and let the teachers run the schools, you have misunderstood me. Educational philosophy calls for quiet contemplation, and education policy should be set by the political process. Teachers' views should be heard along with all the others. Teachers do what they do now in the belief that it is the best they can do within the constraints. It is not good enough. But if we neglect teachers' ideas, we risk losing a valuable counterbalancing force against the simplification of testing.

Reform needs to give a central place to the perceptions of our teachers, because it is only by building on those perceptions that we can elevate the level of teaching. To pursue implicit and explicit standards, teachers work from an intricate conceptualization of education. They tend to see each child as unique, to regard the curriculum as structured by content, and to think of teaching as activity-based. They assess each student's achievement informally according to their own views of educational objectives, and they rely little on information from standardized testing. For the present, teachers do not hold in high esteem the contribution of standardized testing to teach-

This content downloaded from 159.242.252.8 on Thu, 21 Mar 2013 17:16:20 PM
All use subject to JSTOR Terms and Conditions

ing and learning. Our research on teaching in the classroom suggests that it is a mistake to design a revolution in American schools around a national testing program.

1. Research on the automation of education has been summarized by Roy Pea and Elliot Solaway, "The State of the Art in Educational Technology Research and Development: Policy, Issues, and Opportunities," NTIS #OB 88-194 634/AS, Office of Technology Assessment, U.S. Congress, 1987.

2. Many advocates of school reform in America today share a vision of tightened management by central administrators relying on information from student achievement testing that is supposed to indicate the effectiveness of instruction. A countermovement envisions school-based decision making with increased control by teachers. See Norman Fruchter, "Rethinking School Reform," *Human Services Social Policy*, Summer 1989, pp. 16-25.

3. Swedish researcher Ulf Lundgren conceptualized the conditions of instruction monitored by teachers in "Frame Factors and the Teaching Process" (Doctoral dissertation, Gothenburg University, 1972). Seymour Sarason has written cogently on social conditions in the classroom in *The Culture of the School and the Problem of Change* (Boston: Allyn and Bacon, 1971).

4. Dan Lortie, *School Teacher: A Sociological Study* (Chicago: University of Chicago Press, 1975); N. L. Gage, *Teacher Effectiveness and Teacher Education* (Palo Alto: Pacific Books, 1972); Henry Giroux, Teachers as Intellectuals (New York: Bergen & Garvey, 1988); and Barak Rosenshine, *Teaching Behaviors and Student Achievement* (Stockholm: International Association for the Evaluation of Educational Achievement, 1970).

5. Israel Scheffler, *Conditions of Knowledge* (Chicago: Scott, Foresman, 1965).

6. Ted Aoki referred to individualized teaching as "curriculum-as-lived," as opposed to "curriculum-as-plan." See his "Curriculum Implementation as Instrumental Action and as Situational Praxis," in Ted Aoki et al., eds., *Understanding Situational Meanings of Curriculum In-Service Acts: Implementing, Consulting, Inservicing* (Edmonton: University of Alberta, Curriculum Praxis Monograph Series 9, 1983), pp. 3-17.

7. Those calling for more emphasis on a "core" curriculum have included Gary Fenstermacher and John Goodlad, eds., *Individual Differences and the Common Curriculum: 82nd NSSE Yearbook, Part I* (Chicago: National Society for the Study of Education, University of Chicago Press, 1983).

8. Douglas Archbald and Fred M. Newmann, *Beyond Standardized Testing: Assessing Authentic Academic Achievement in the Secondary School* (Reston, Va.: National Association of Secondary School Principals, 1988).

9. Carol Tittle, Kathy Kelly-Benjamin, and Joanne Sacks, "The Construction of Validity and Effects of Large-Scale Assessments in the Schools," in Robert E. Stake and Rita G. O'Sullivan, eds., *Advances in Program Evaluation, 1B: Effects of Mandated Assessment on Teaching* (Greenwich, Conn.: JAI Press, 1991), pp. 233-54; David A. Goslin, *Teachers and Testing* (New York: Russell Sage Foundation, 1967); Martin Hotvedt, "A Case Study of Standardized Test Use in the Public School" (Doctoral dissertation, University of Illinois, 1974); and Joan Herman and Donald Dorre-Bremme, "Uses of Testing in the Schools: A National Profile," *New Directions for Testing and Measurement*, vol. 19, 1983, pp. 7-17.

10. It could be that teachers are failing to use a useful resource and need coaching. Leslie McLean of the Ontario Institute for the Study of Education finds it more reasonable to conclude that test data do not fit the experiential understandings of teachers. See Leslie McLean, "Achievement Testing – Yes! Achievement Tests – No," *E + M Newsletter*, OISE Educational Evaluation Centre, 1982, pp. 1-2; and Eleanor Dale Costello, "Kaleidoscope Patterns: Art Education in an Elementary Classroom" (Master's thesis, University of British Columbia, 1988).

11. Robert E. Stake, *Validity and Invalidity of Standardized Mathematics Achievement Testing* (Madison, Wis.: National Center for Research in Mathematical Sciences Education, forthcoming).

12. Reported in Bernadine Stake, Robert Stake, Laura Morgan, and James Pearsol, "Evaluation of the National Sex Equity Demonstration Project, Final Report, 1980-1983," Center for Instructional Research and Curriculum Evaluation, University of Illinois, 1983, p. 221.

13. One of the finest efforts toward creating a science of education occurred in the 1970s when Lee Cronbach and Richard Snow of Stanford University attempted to pin down relationships between aptitudes (as indicated by tests) and pedagogical strategies. For example, did certain children learn better through practical application while others got more from abstract explanations? See Lee Cronbach, *Aptitudes and Instructional Methods: A Handbook for Research on Interactions* (San Francisco: Jossey-Bass, 1977).

14. Surely some educators are so experienced with testing and curriculum development that they can act as consultants to guide education reform from test scores? An Australian researcher spent more than a year in the Midwest, interviewing and observing in many districts, trying to find one such sage. He concluded that such people do not exist. See Norman Bowman, "A Search for Instances of District Use of Aggregated Test Data for Curriculum Improvement" (Doctoral dissertation, University of Illinois, 1979).

15. Neil Sheehan, *A Bright and Shining Lie: John Vann and America in Vietnam* (New York: Random House, 1988).

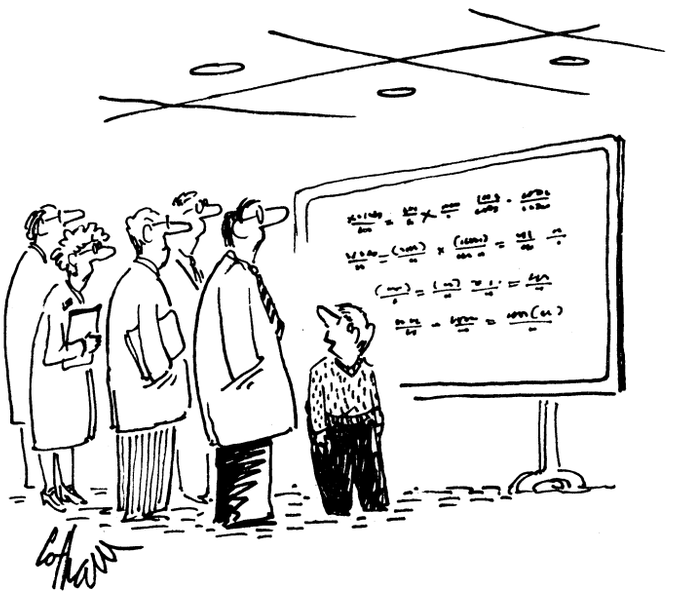
16. David Halberstam, *The Reckoning* (New York: Morrow, 1986).

17. Thomas J. Peters and Robert H. Waterman, Jr., *In Search of Excellence: Lessons from America's Best-Run Corporations* (New York: Warner, 1984).

18. Walter Haney and George Madaus, "Effects of Standardized Testing and the Future of the National Assessment of Educational Progress," working paper prepared for the NAEP Study Group, Center for the Study of Testing, Evaluation, and Educational Policy, Boston College, Chestnut Hill, Mass.

19. We should not be reluctant to look at how educational authorities around the world are reforming education. The Swedish Parliament dismissed the 800-person National Board of Education and replaced it with an agency for assisting teachers. The Ministry of Education of Victoria, Australia, has created a system of teacher assessments of standardized projects and portfolios to decide who will go to college. The United Kingdom has piloted "standard assessment tasks," but teachers find the load excessive. Ontario continues to revise its curriculum along lines supported by teacher unions, without reliance on state or federal testing. Some ministries seek to draw from science and technology without undermining existing pedagogical arts; others do not.

20. Andrew Porter, "External Standards and Good Teaching: The Pros and Cons of Telling Teachers What to Do," *Educational Evaluation and Policy Analysis*, vol. 11, 1989, p. 354. See also Harry Torrance, "Evaluating SATs – the 1990 Pilot," *Cambridge Journal of Education*, in press. K



"Gee, dad, it sure is nice of you and your research staff to help me with my homework."

This content downloaded from 159.242.252.8 on Thu, 21 Mar 2013 17:16:20 PM
All use subject to JSTOR Terms and Conditions