DOCUMENT RESUME

ED 392 336 TM 024 606

AUTHOR Spencer, Thelma L.
TITLE The Testing Mania: Wreckage or Reform?
PUB DATE 19 Apr 95
NOTE 16p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (San Francisco, CA, April 19-21, 1995).
PUB TYPE Information Analyses (070) -- Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Accountability; Curriculum; *Educational Assessment; Educational Change; Educational Improvement; Elementary Secondary Education; *National Competency Tests; Performance Based Assessment; Standardized Tests; Test Construction; *Testing; *Test Use
IDENTIFIERS *Alternative Assessment; America 2000; National Education Goals 1990; *Reform Efforts

ABSTRACT

This paper synthesizes information derived from five sources regarding issues in testing as they related to the American 2000; an education strategy and the national education goals. The focus is on proposed national tests. The issue of national testing extends far beyond testing itself because it touches on the structure of the U.S. educational system. Many educators argue that negative consequences are almost inevitable in a national testing program and that curriculum and goals will be manipulated to promote higher test scores. Even advocates of a national test recognize that accountability poses a serious problem and that curriculum as it exists and test content often do not complement each other. Overall, mainstream educators appear to have accepted the inevitability of some kind of national test, but most doubt that it will result in substantial educational improvement. Schools across the country are overwhelmed now by the standardized tests in use, and they are beginning to attempt to use alternative assessments to measure student achievement. If the proposed American Achievement Tests do not reflect the new paradigm of teaching students to think critically and solve problems creatively, then they should be redesigned. (Contains five references.) (SLD)

THE TESTING MANIA; WRECKAGE OR REFORM?

by

Thelma L. Spencer, Ed.D.

Symposium
"Measuring the National Goals for Minority Students"
presented at the Annual Meeting of the
National Council on Measurement in Education
San Francisco, California
April 19, 1995

THE TESTING MANIA; WRECKAGE OR REFORM?

This paper synthesizes information derived from several sources deemed to be sufficiently inclusive and recent to permit one to arrive at a tentative answer to the question posed in the topic of the paper. To that end, I gathered the following materials:

- **Educational Leadership**, February 1995
  Theme of the issue - School Reform: What We've Learned
- **Educational Leadership**, March 1995
  Theme of the issue - Aiming for Higher Standards
- Educational Testing Service **1994 Annual Report**
- **Voices From The Field: 30 Expert Opinions on America 2000, 1991**
- Executive Summary of **The National Education Goals Report, 1992**

FINDINGS

Each of these publications provided a unique perspective on the testing issue as it relates to **America 2000: An Education Strategy**, and the nation's efforts to achieve the national education goals.

While the GAO investigators reported in the Executive Report that U.S. students do not seem to have been overtested, educators appear to be about evenly divided on the merits of increased testing of students as a means of reforming American education. Among those who have spoken out on the subject we find such respected and knowledgable authorities as Linda Darling-Hammond, Jeannie Oakes, Marshall Smith, Gerald Bracey, Lawrence Stedman, and Gary Orfield.

None of these individuals can be described as being opposed to reform of our current educational system, nor are any of them opposed to evaluation of the results of the reforms that have been or may be implemented. I will comment briefly about them, as well as others in the remarks that follow.

The GAO and the Money

In 1991, the GAO reports, this country began debating in earnest the proposition that the United States should adopt a national examination system. "It soon became apparent that the debate lacked some key information about the present extent and cost of testing, as well as the likely cost of a national examination system"

The investigators found, the report says, systemwide testing (that is, the testing that is given to all students at any one grade level in a school district) took up about 7 hours for an average student in 1990-91 (half of that time in direct testing and half in related activity), and cost about $15 per student (including the cost of the test and staff time). They estimated that such testing cost about $516 million nationwide that year, and that a national examination -- depending on whether it is based on multiple choice or performance testing -- would cost, respectively, about $160 million or about $330 million annually.

Most of the debate on expanded testing has centered on major issues of what to test, how to test, and how to use the results. Not much attention has been given to date to the question of how much and what kind of testing there is now. Yet the likely success of future testing may be related to the size, nature, and cost of current efforts about which there exist only wide-ranging, conflicting, and highly uncertain estimates. These range from 30 million to over 127 million standardized tests administered per year, at a cost of from $100 million to $915 million.

The congressionally mandated National Council on Educational Standards (NCEST) declined to provide a cost estimate in its report recommending a national testing system, and others' estimates have ranged from a few million dollars a year up to $3 billion -- that's billion with a bee!

It would not be an overstatement to say that among these researchers, testing experts and educators there is no general opposition to testing as such. It would also be safe to say that each has, in one way or another expressed concerns over the purpose, quality, and locus of control over the content and administration of further tests.

Some of the voices in the field do not even get to this stage of the debate since they find **America 2000** flawed at the outset, and question its very valdity as a concept to be given serious consideration.

It follows, then, that the issue of national testing extends beyond testing itself. Some wonder aloud if the proposed American Achievement Tests system is a covert attempt to establish a national curriculum. As a top-down initiative based on existing, primarily multiple-choice, testing technologies, the AAT would have education regressing from the current innovations found in school districts across the country. Such local efforts could be seriously undermined if a system pegged to national norms were instituted.

The National Education Goals as promulgated in President Bush's **America 2000** pushes the nation even farther along the road to universal testing in American schools. Goal 3: Student Achievement and Citizenship reads as follows:

> By the year 2000, all students will leave grades 4, 8, and 12 having demonstrated competency over challenging subject matter including English, mathematics, science, foreign languages, civics and government, economics, art, history, and geography, and every school in America will ensure that all students learn to use their minds well, so they may be prepared for responsible citizenship, further learnings, and productive employment in our Nation's modern economy.

Of the accompanying six objectives, we are concerned with two that focus on evaluating the outcomes of these efforts. It became clear how this evaluation was to be carried out when the President proposed the development and establishment of an American Achievement Test (AAT). The pertinent objectives are:

- The academic performance of all students at the elementary and secondary level will increase significantly in every quartile, and the distribution of minority students in each quartile will more clearly reflect the student population as a whole.

- The percentage of all students who demonstrate the ability to reason, solve problems, apply knowledge, and write and communicate effectively will increase substantially.

The first shove in the direction of universal testing came in the early 1980s when the spate of reports on the quality of the education provided to American children and youth seemed agreed that the best description was that it was "mediocre" at best. The reports raised levels of concern and even created some anxiety among citizens responsible for the nation's schools--local and state board members, state and federal law-makers, teachers and teacher educators, and, not least of all, parents and taxpayers. The Charlottesville meeting provided the next impetus for the advocates of reform.

While each group had its own agenda, and offered a solution of one kind or another, there appeared to be general agreement that the quality of efforts to reform and improve the educational

enterprise could best be evaluated by using standardized tests. The proposal to create the AAT became the subject of speeches, papers at professional meetings, bones of contention between and among the vested interest groups, including test developers, and a magnet for considerable ill-informed and misinformed debate.

Who says what about a national test

In **Voices From the Field**, a compilation of 30 expert opinions on America 2000, Linda Darling-Hammond writes that "standardized tests [are] increasingly employed as measures of student achievement and as arbiters of decisions about student placements, teacher competence, and school quality. Some policy makers have also sought to use tests to 'hold schools accountable' by triggering rewards, sanctions, or remedial actions." (p. 15).

She has written and spoken at length about the proposal to extend state and local efforts to the national level through the establishment of a universal testing system. This may not be an idea whose time has come in Darling-Hammond's view. She avers that available evidence suggests that, by and large, current testing policies have not had many of the positive effects that were intended for them, while they clearly have had many negative consequences for the quality of American schooling and for the equitable allocation of school opportunities. (p. 15). She writes:

> [The] shortcomings of American testss have become more problematic as test scores have been used to make important educational decisions. As schools have begun to "teach to the tests," the scores have become ever-poorer assessments of students' overall abilities. (p. 15).

Another of the expert voices, Jeannie Oakes, contends that negative consequences are almost inevitable for substantial proportions of the student population, particularly when students are required to prove their competence in high-stakes testing and certification programs.

These negative consequences are exacerbated by two features of schooling in America about which policy makers have been, in Oakes' words, "curiously silent." She cites the uneven distribution of resources and opportunities, resulting nearly always in less of the pot for disadvantaged and minority students. She points out, as well, the propensity among those who control the purse strings to use test results to close off rather than open up opportunities for these same students.

In a similar vein, Marshall Smith asserts that a national test for students is an obvious elixir for politicians. He says that it fits the American political habit of looking for an inexpensive, quick fix. The AAT would be given in the 4th, 8th and 12th grades in five core academic areas. It would embody new standards for school performance and monitor the country's progress toward achievement of the National Education Goals. What more could one ask for?

What Smith calls the "national test rocket" was not launched in a vacuum. It was powered by the fuel of poor international comparisons on achievement tests and by business sector claims that high school graduates are ill-prepared for the work world. (**Voices From The Field**, p. 21). It was suggested that periodic status reports, maybe even a "national test," based on some vaguely defined national standards would save us from future international embarrassment.

Smith asks, "A national test -- to what end?" He describes the perceived functions of such a system as serving three major purposes:

1) National tests could monitor the progress of student achievement for the country as a whole, for states and districts and even for schools and classrooms.

2) National tests could be used to inform decisions about the educational future of individual students. Tests used in this regard lead to decisions that help determine the life-chances of students. These decisions require individual student scores and include placement in secondary school tracks, college admissions, and even employment opportunities.

3) National tests could be used to hold institutions accountable for achievement of some set of standards.

Accountability poses a serious problem for both proponents and opponents of an AAT. One major reason for the failure of the nation's present system of educational accountability, Smith contends, is the lack of congruence between the content of the tests being used and the content of the curriculum that students are taught. (p.22).

He maintains that if we are to succeed in improving student achievement, then we must change the nature of teaching and learning. He says it is not sufficient to alter the test and curriculum specifications. New curriculum materials and well-prepared teachers are a prerequisite. He writes: 'The central implication of a seriously different curriculum and challenging new standards . . . implies a systematic approach to reform." (p. 22).

Gerald Bracey and Lawrence Stedman engaged in something of a verbal brawl in the ASCD Journal, **Educational Leadership, in February/March of this year**. Stedman took the position that the threat to public education is not imminent collapse. He says the real issue is complacency that all is well. A balanced appraisal, he argues, shows that major structural changes are needed. He charges that Bracey, Carson, et al., Huelskamp and Rotberg have engaged in creating a myth about the condition of of American education.

Following his evaluation of the same performance data that these researchers used, Stedman concluded that American students have done well internationally in reading, but their performance has been dismal in mathematics, uneven in geography, and often poor in science, particularly among the 17-year-olds.

He claims that the new mythology ignores the large body of evidence that shows that general education performance is low and has been for decades. According to Stedman, this evidence supports the position that fundamental school reforms are warranted -- including far-reaching overhauls in funding, governance, organization, pedagogy, and evaluation.

He advises educators and school district officials to ask the following questions:

- How well are the students learning the curriculum?
- Can students demonstrate high levels of general knowledge as well as sustained problem solving?

To judge this and encourage high intellectual standards, the school district should consider adopting performance assessment, standardized tests of subject matter knowledge, and exhibitions of mastery for graduation.

In sum, he says, while education at the national level is not deteriorating, it is certainly not succeeding. At the local level, concerned educators and community members should determine how extensively schools need to be reformed and in which direction to take them. (**Educational Leadership**, February 1995, p. 84).

Bracey denies that he is complacent, and says that neither are any of the others cited by Stedman guilty as charged. He says that the jury is still out on whether American schools have failed, but conflicting evidence and cause for optimism are not complacency. "We simply do not know," Bracey writes, "how our educational system is functioning, save for anecdote and impression." (p. 76).

Using the same data, Bracey and Stedman reach opposing positions on the myths about the status of education in the United States:

Myth 1: There really wasn't an SAT decline. Stedman says there was. Bracey says that his statement there was "no decline" was only half in error because there was a gain in math but a decline in verbal.

Myth 2: Test scores are at an all-time high. Stedman says the claim is only partly supported by the evidence; Bracey says some scores are at all-time highs.

Myth 3: Our top half is internationally competitive . Stedman says only the top 10 percent are competitive, using Tanner's 1993 narrow claim about performance in science. Bracey, on the other hand, uses the same data to support his position that it makes little sense to speak of a nation's relative standings based on average test scores because the variability within any nation is huge.

Myth 4: The crisis in education is not general; it concerns inner-city schools and poverty. Stedman alleges that Bracey, et al explicitly reject the view that the "typical school is failing," and that reform should focus on urban minority students because the system works well for others. Stedman argues the new mythology ignores the vast

> body of evidence that shows that general educational performance is low and has been for decades. Bracey, on his part, points out that rural poverty is even more dire than its urban counterpart. (a weak rejoinder in my view)

Gary Orfield, yet another of the expert voices from the field offers this acerbic comment on the current goal-based reform movement:

> It has been so long since we have had a serious proposal for a major change in federal education policy that we have come to treat public relations strategies as if they were coherent policies. Old ideas are treated as if they were new. (p. 7).

He charges that America 2000 is not a plan for American education but a plan for re-electing the President. When he says that the effects of the plan on the schools will be exceedingly small, he leaves little doubt about his sentiments regarding the current reform movement. Neither does he hedge on where he stands with respect to the testing issue as illustrated by this lengthy quotation from his paper entitled, "Choice, Testing and the Re-Election of a President" in **Voices From The Field**.

> Testing proliferated during the 1980s in virtually all states, but had little effect on achievement levels. A strong re-emphasis on standardized tests began in the late 1970s, a central recommendation of the Nation at Risk report. Now the President wants a new set of national tests. The 1980s tests were often used to raise standards by flunking children who could not meet certain test scores, a policy that has failed to produce educational gains, increased spending for repeating the same grades of school, and raised dropout rates. Problems with the new national test system include the danger of intensifying curriculum domination by tests, the costs in dollars and school time of more tests, the possibility of bias, and the probability that test scores will be used in misleading ways for inappropriate comparisons of schools and districts. Although these problems may be surmountable, there is little to suggest large benefits from another layer of testing. (p. 7).

He obviously has serious doubts that more widespread use of tests or the establishment of a national testing system can transform American schools by the year 2000.

There are, of course, two sides to every question, and diametrically opposed positions may be equally valid. Weighing in as proponents of testing are the developers of standardized tests themselves. One of these is the Educational Testing Service (ETS). ETS is the headquarters of the National Assessment of Educational Progress, a nation-wide testing program funded by the National Center for Education Statistics.

ETS sees its mission as lending support, through providing leadership on test-related educational issues to those confronted with what its calls the inevitable linking of education and testing in achieving the desired quality of education. The **ETS 1994 Annual Report** states that:

> Testing has critical roles to play by reflecting effectively high learning goals [i.e., the ability to think critically, to solve new types of problems] and by communicating them to a public that has only a vague understanding of educational rhetoric about higher-order thinking skills. Tests provide concrete examples of the types of tasks learners of higher-order skills should accomplish. (emphasis added) These concrete examples serve to sharpen our understanding of the goals themselves and provide a means to communicate these goals more effectively. (ETS, 1994)

An AAT: Yes or No?

It is not surprising to find that there is not a great deal of enthusiasm among the voices in the field nor even among the troops in the trenches for the development and implementation of a national achievement test system. It is also interesting to find that neither is there a great deal of opposition to the idea. Mainstream educators, for the most part, appear to have accepted the inevitability of the AAT, but all of them urge a careful look at the implications to be drawn from using the test data. Of the educators whose positions you have already heard, there are several others I would call to your attention.

Michael Timpane of Teachers College, for example, urges caution in the movement toward national tests. He says that even as the time has come for the nation to commit itself to a set of national education goals and to high standards of school achievement needed to accomplish them, we **must** move forward to define these goals before we define any system of national tests.

> "The tests do not fully track what is taught, and they will distort and even dictate the standards. We know that the

> tests we have today are not adequate to measure progress toward the kinds of goals now being considered. We also know that too often the significance of the tests is exaggerated -- instructional patterns are changed so that students will score higher on them, and the results are misused as a means of labeling and tracking children who score poorly. And we know, finally, that awarding funds on the basis of test results creates perverse and dangerous incentives. Expanded national assessments may be necessary, even desirable, but they are not sufficient to encourage and ensure greater learning, and they could be harmful if hastily imposed. **(Voices From The Field, p. 19)**

James Rosenbaum, Professor of Sociology at Northwestern University, charges that **America 2000** poses bold goals for American education, but it lacks sufficiently bold and effective strategies to achieve them. The designers proposed national achievement tests as a means to ensure that all schools meet acceptable standards. "Weak schools cannot be made stronger by tests alone, and good tests are a decade away. Even then, we may not know how to use them. (emphasis added) (Voices, p. 15)

Joan Wills, Director of the Center for Workforce Development, is not surprised that much of the initial response t the proposed agenda of the National Goals Panel and the four-point education strategy focused on the political consequences of a national testing system for the schools. This apparently only served to confirm her already dim view of what she characterizes as the politically-negotiated voluntary testing system advocated by **America 2000**.

William Wayson of The Ohio State University writes that **America 2000** is fatally flawed, and he appears to be especially incensed by the use of the four-train metaphor and the allusion to the ticket that says we are bound for 2000. In his words, "no track has been built to get there." (Voices, p. 48)

With respect to testing, Wayson says America 2000 is too narrow a view of the school's role and of what good teachers do. It adheres to the limited purposes that can be tested by norm-referenced tests. Education must go beyond what can be tested; it must improve the quality of life. He goes on to say that what can be tested easily and efficiently is a small part of what the fully educated person is and falls short of what the technological world requires to keep America on top.

Wayson tells us that past experience suggests that proficiency tests will focus on isolated facts rather than

integration, synthesis, analysis, application, and understanding. Test scores can be raised without any child's learning anything, and the public may be lulled into thinking that the problems are solved. The extreme difficulty of eliminating test bias is dismissed, he writes, with the "fatuous assertion that tests will be screened to eliminate it." (p. 49)

Jose Cardenas of the Intercultural Development Research Association, tells us that comparisons of schools, districts and states will demonstrate only what is already known: that some segments of the population perform deplorably in the education process. School, district and state characteristics will be reflected in these comparisons. . . . The challenge is to provide assistance to students not performing well, not to further stigmatize them. (p. 28)

Cardenas observes pessimistically that it is doubtful that minority, disadvantaged, limited English proficient, immigrant and migrant students will do any better or worse (emphasis added) on a new national exam than they are doing on the standardized national and state examinations currently in use. (p. 28).

Harold Howe II, former U.S. Commissioner of Education and Professor or Education at the Harvard Graduate School of Education, perceives America 2000 to be a flawed agenda for meeting the education needs of American children and youth. He says that it is not clear to him what is planned or wanted with regard to national testing.

Certainly we need new and better tests that do more to help teachers and students in the learning process than do current standardized tests, he maintains. But such tests are more useful in the classroom than they are for purposes of establishing comparisons.

> If what is intended is making new and better tests available nationwide to those in schools who want to use them, I say "Amen." But if the plan is to spring new tests on kids as a way of forcing teachers to adapt to standards, I'm dead against it. (p. 27)

The final voice you should hear is that of Bernard Watson, who has worn many hats during his distinguished career in education: public school teacher and principal; Deputy School Superintendent in Philadelphia, professor and Academic Vice President of Temple University; President and CEO of the William

Penn Foundation. Watson thinks all Americans should support the National Education Goals. He points out that the designers of America 2000 recognized that assessment is central to knowing whether students are achieving the academic gains the nation expects of them.

But this is not enough, he says. It is equally important to note that testing should come not only while schools are in the **process** of reformation but **after** they have been reformed. As others have pointed out, we already hasve a pretty good idea of which schools and school systems are doing well and which are not, looking at data from tests now in use. This does not, however, preclude the need for the development of new and better modes of assessment to characterize education reform. (p. 32)

And, finally --

We have heard the voices from the field urging us to go slowly, to take our time, to be sure that the reforms we want and must have are what we want and are achievable. We are still so early in the process that we are not yet ready to put in place anything as ambitious or as superficially conceptualized as the proposed American Achievement Tests system.

Perhaps these words from the November 1994 Washington Post Education Review will tell you how it is with the troops in the trenches whom the President and his advisors seem to have forgotten or just plain ignored.

> Each fall, when standardized test scores are released, teachers hold their breath. Those numbers, used by state officials to measure how effective educational institutions are and whether students are learning, can determine the fate of a school district, a principal and, especially, a teacher. ("Will That Be on The Test?")

Schools across the country are overwhelmed right now by standardized tests. In some districts, students are required to take as many as 12 tests a year. Some school districts have begun moving away from the testing programs that have been used nationally for years to gauge how well teachers are doing their jobs. In place of norm referenced standardized tests these districts have begun to use alternative assessments, tests that measure a student's ability to perform tasks, including essays, open-ended problem solving in math and science and in some districts requiring students to compile portfolios of what the students consider their best work.

If the proposed American Achievement Tests do not reflect the new paradigm of teaching students to think critically and solve problems creatively, then the designers need to go back to the drawing board. If they do not, then, like Harold Howe, I'm dead against it.

NCME, 1995
Thelma L. Spencer, Ed.D.

REFERENCES

Bracey, Gerald W. (March 1995). "Stedman's Myths Miss the Mark." **Educational Leadership** 52: 75-80.

National Education Goals Panel (1992). **Executive Summary of the National Education Goals Report.**

National Education Goals Panel (1994). **Building a Nation of Learners 1994.**

Stedman, Lawrence (February 1995). "The New Mythology About the Status of U.S. Schools." Educational Leadership 52: 80-85.

W.T. Grant Foundation Commission on Work, Family and Citizenship (1991). **Voices From The Field: 30 Expert Opinions on America 2000, The Bush Administration Strategy to "Reinvent" America's Schools.**