DOCUMENT RESUME

ED 353 272 TM 018 979

AUTHOR Hogan, Thomas P.
TITLE Prospects and Problems for a National Test: Some Reflections of a Test Author.
PUB DATE Apr 92
NOTE 12p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (San Francisco, CA, April 21-23, 1992).
PUB TYPE Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Cost Effectiveness; Curriculum Development; Educational Assessment; *Educational Improvement; Educational Objectives; Elementary Secondary Education; Futures (of Society); *National Competency Tests; National Programs; Psychometrics; *Student Evaluation; Test Construction; *Testing Problems; Testing Programs; Test Use

ABSTRACT

Reflections on the proposal for creation and implementation of a national test are presented from the perspective of a test author. The most readily identified characteristic of the proposed national test is the intensity of debate surrounding it. Another easily identified characteristic is the anticipated effect. While proponents expect higher levels of academic achievement, and opponents fear reduced achievement, it seems probable that there will actually be no perceptible effect on achievement levels or on anything else, unless it is the national debt. The matter of curriculum is not a significant impediment to the development of a national test. In effect, a "de facto" national curriculum already exists, and a national test is likely only to damage the perception of local curricula. The technical and psychometric problems associated with a national test can probably all be solved. The most under-rated, or under-discussed, problems are the logistical ones of testing, scoring, and result reporting. There is no evidence that the proposed test intends to acknowledge individual differences among test takers, and that is a major problem. While there seem to be no compelling reasons why a national test cannot be built, it will not accomplish what is hoped for it. It will be expensive and will not sustain itself because of its failure to deal with individual differences and its ill-defined purpose. (SLD)

PROSPECTS AND PROBLEMS FOR A NATIONAL TEST: SOME REFLECTIONS OF A TEST AUTHOR

Thomas P. Hogan
University of Scranton

PROSPECTS AND PROBLEMS FOR A NATIONAL TEST: SOME REFLECTIONS OF A TEST AUTHOR

Thomas P. Hogan
University of Scranton

My understanding of the task set for me by Dr. Rudman, our chair, is to reflect on the proposal for creation and implementation of a "national test" from perspectives developed in my role as a test author. Let me begin by identifying briefly what is meant by the terms a "national test" and my role as "test author."

By a "national test" is meant that creature envisaged in the America 2000 proposal currently espoused by the executive branch of the federal government. One must admit that reflecting on this proposed national test presents the proverbial moving target problem, since the proposal sketches some broad outlines for the test, and makes a number of specific assertions, but one does not really have a fully developed, detailed proposal to examine.

By my role as a "test author" I refer to about 25 years of experience in developing nationally standardized tests which have enjoyed reasonably widespread use in elementary and secondary schools. The experience includes conceptualizing the tests, writing test exercises, trying them out, standardizing the tests, writing interpretive materials, and working with school systems and states in their utilization.

==========

Paper presented as part of the invited symposium on "National Goals and National Testing" at the annual meeting of the National Council on Measurement in Education, San Francisco, April, 1992.

While reflecting on the proposal for a national test, one is struck first and most vividly not by any characteristic of the proposed test but by the ferocity of the debate surrounding it. On the one hand, there are fanatic opponents, who see the proposal as philosophically unsound, technically impossible, with cataclysmic consequences, and all together evil incarnate. On the other hand, there is a significant group of rabid proponents, who see the proposal as accomplishable with relative ease and the simple, convincing answer to most of the problems facing education today. There appear to be not many people in between these polarized extremes, at least not many whose voices are being heard. I note in particular the relatively few comments being made by persons with significant experience in actually creating and producing nationally used tests of the general character contemplated for the proposed national test: A curious situation, to say the least.

Shall I, then, step in where wiser heads have chosen not to venture? Well, yes, but only at Dr. Rudman's urgent request -- and with foreknowledge that it were probably best not to do so. But, having agreed to do so, let's have at it. What do I make of this proposed national test, from my perspective as a test author? Without pretending to exhaust the topic, let me comment on five specific areas relevant to the proposal.

1. Anticipated Results. As noted above, the most readily identified characteristic of the proposed national test is the intensity of debate surrounding it. Perhaps the second most

easily identified characteristic of the proposal is the anticipated outcome of having a national testing program. The proponents of the program anticipate -- and a great expectation it is -- that the tests will focus attention on the educational enterprise and thereby lead to (dramatically) higher levels of achievement. The opponents of the proposed national testing program expect, more or less the opposite, or worse. They expect that achievement levels will decline, the curriculum will stagnate, and self-esteem will (probably) be damaged.

What is my personal expectation about the probable impact of a national testing program on achievement levels? My personal expectation is that a national testing program would have no perceptible effect on achievement levels in one direction or the other. Opponents of testing cite the correlation between increases in state-wide testing program and a decline in achievement levels -- roughly in the later 1970's and throughout the 1980's -- as evidence that testing leads to lowered achievement. They fail to point out that during the period of most rapid expansion in standardized testing, roughly from the late 1940's to the mid-1970's, achievement levels rose rather continuously. But for none of these trends is there any evidence of a causal connection. In the vast array of influences on the lives of students and teachers, most of which occur outside of the immediate school context, the presence or absence of a standardized test has, I have concluded, relatively little influence. And, I would anticipate that the introduction of yet

another standardized test, this time The National Test, would similarly have no perceptible effect on achievement levels of American's youth, nor on anything else, except perhaps the national debt.

At root, the problem here is one of defining the purpose of the test. In my experience as a test author or test development project manager, the very first step in the process is to fix clearly the purpose of the test. Much else follows from this purpose. As far as I can determine from the America 2000 proposal and associated documents, the principal -- perhaps sole -- purpose of the proposed national test is to serve as a motivator: To motivate students, teachers, administrators, parents, legislators, and the public at large. This is probably the most peculiar purpose I have ever encountered for a proposed test. Quite apart from this peculiarity, the problem is that the stated purpose does not give much (if any) guidance for subsequent stages of test development.

2. A National Curriculum? Let me turn now to the question of whether the National Test presupposes (now) or will dictate (in the future) a national curriculum. Interestingly, the America 2000 document (p. 32) addresses this question directly: "Do national tests mean a national curriculum?" And, amazingly, the response is "No." This is sheer, unadulterated nonsense. As soon as one says that the National Tests will cover English, mathematics, science, history and geography, (and not, for example, music, Latin, theology, and military training -- all of

which, by the way, have ample historical precedent for inclusion in the school curriculum) one has posited a great deal about what the curriculum is or should be.

Further, anyone who has put his or her hand to the task of preparing an achievement test knows that one begins with an outline of content -- a test blueprint -- which reflects an outline of what is covered in the curriculum area to be tested. In fact, the America 2000 document (p. 45), after asserting that the tests do not presuppose a national curriculum, goes on to specify, in broad terms, exactly how that curriculum (and, therefore, the test content) will be defined. The answer is: essentially, by professional educational associations, which have been engaged in such business for roughly as long as they have existed, and which have served as one of the sources of test blueprints for standardized achievement tests for as long as such tests have existed.

One wonders why the America 2000 authors would make such a silly assertion as that a national test does not presuppose (or impose) a national curriculum. I suspect it is to avoid clashing with the firmly entrenched notion of local autonomy regarding the school curriculum. Every school I have every worked with believes it has a unique curriculum. Never mind the fact that it is virtually indistinguishable from the curricula of 90% of the other school districts in the nation. The belief is still strongly held and jealously guarded.

Regarding this matter of a national curriculum, let me make

two observations based on my experience as a test author. First, at least in the areas where I have most frequently worked on test development, i.e., in reading, mathematics, and English, I concluded long ago that we have a de facto national curriculum. The similarities among various basal textbooks, local districts' statements of objectives, state curriculum guides, and other such documents are far greater than are their differences. The opportunity exists for local variation (and that may be an important point) but it is not often exercised.

Second, what about the concern that a national curriculum will lead to stagnation or the frustration of reform? Part of the national curriculum seems to be that it is subject to periodic, sweeping reform, quite independent of what is contained in national tests. Proponents of curriculum reform have long told school administrators and teachers that existing tests are not valid for their novel objectives; and the claim has been readily accepted as the reforms proceed. I am reminded of the modern math movement of the 1960's, the linguistics movement of the 1970's, the whole language movement of the 1980's. None of these movements seemed the least restrained by the content of existing tests.

From the foregoing observations, my conclusion is that the matter of curriculum is not a significant impediment to the development of the National Test (except possibly in tampering with the perception of local control of the curriculum and I must confess to being uncertain about the importance of this

variable).

3. Technical, Psychometric Problems. A host of questions have been raised about the technical or psychometric feasibility of developing and implementing a national test. What do I make of these? Frankly, I don't view these problems as very significant. The problems have all been faced before. Their solutions, while not perfect, have been applied repeatedly in an assortment of testing contexts. They will be applied again, much as they are routinely applied now (although sometimes with the claim that they are imaginative solutions to heretofore intractable problems -- such a claim being made for purely political purposes).

A somewhat separate technical issue relates to the promise that the new National Test will incorporate "authentic" assessments (good) as opposed to relying exclusively on traditional, objective measures (bad). Of course, these authentic assessments are not available for public scrutiny and professional examination. They are simply accepted as a yet-to-be delivered good. It is too kind to say that perhaps the promise will materialize. The fact is that such assessments have been tried over and over again, only to be found wanting in terms of the most fundamental criteria of measurement. Although the critics of "objective" testing like to claim that such tests were developed to capitalize on the efficiencies of machine-scoring, that is not the case. "Objective" tests were developed,

historically, because they were found to be both more reliable and more valid than the woefully inadequate testing practices of the past, many of which would today be termed "authentic." (And, it needs to be noted, machines to score these "objective" tests were developed a number of decades after the objective tests had established their superiority.)

Let me conclude these comments on technical matters with this note of anticipation. Assuming that the National Test is forthcoming, I can hardly contain my excitement about discovering what percentile points, for the various grade levels and subject areas, will be used to define "world class standards." After all, if I believe what I read about what is to be done with the tests, it is inevitable that such definitions must be given -- and I'm dying to learn what they are.

4. Logistics. Perhaps the most under-rated -- or more exactly, undiscussed -- set of difficulties to be faced by the proposed National Test relates to logistics. What is proposed is not just the development of a test, but the implementation of an entire system, including distribution of materials, administration, collection, scoring, and reporting of test results. And reports are not simply made about aggregated results but about individual students to parents, colleges, and employers. My experience as a test author tells me that there are infinitely more problems to be solved with these logistical matters than with just the development of the test.

If indeed this is to be a high-stakes test -- one which is used by employers in hiring decisions, colleges in admissions decisions, legislators in funding decisions, schools in promotion/graduation decisions -- then such problems as test security and score reporting mechanisms, when viewed in the context of a nationwide program, are simply enormous. And the costs will be staggering.

5. Individual Differences. One of the most obvious facts in the world of testing is that of individual differences: very large differences. There are some second graders who write more fluently than the typical eighth grader; some eighth graders who write no better than the typical second grader. There are sixth graders who are more proficient at mathematics than the typical college senior; some college seniors who are no more proficient than the typical sixth grader. In some ways, such differences are at the root of what we know today as educational and psychological measurement.

I see no evidence that the proposed national test intends to acknowledge these individual differences. In fact, the differences seem to be an embarrassment, studiously avoided. I do not believe one can build systems which sustain themselves over long periods of time while avoiding such large facts.

Let me conclude. Based on my experience in building tests, from a technical viewpoint I do not see any reason why it is not

possible to build a national test. We have the psychometric expertise to do a creditable job; and there is certainly a curriculum base to use as a springboard.

However, if the test is built, it will not accomplish the purpose set out for it. Its implementation will be savaged by logistical problems; its costs will be enormous. And it will not sustain itself over any significant period of time because of its failure to deal with the fundamental issue of individual differences, as well as because of its ill-defined purpose.