



Curricular Accountability Through Testing¹

Ted Ward

From the point of view of a curriculum analyst, testing is a powerful tool. Testing is a shaper of the curriculum. Testing is second only to instructional materials as an index of what the teacher will be doing in the classroom. Testing constitutes a set of red flags and green flags, giving the teacher and the student bit-by-bit feedback about what is *really* important in the educational establishment.

Tests should be a *reflection* of what is important, rather than a *determiner* of what is important for boys and girls to learn. When either tests or instructional materials are out of tune with what children most need to learn, the school may be doing the wrong thing—and perhaps doing the wrong thing rather effectively! When instructional programs are not carefully related to the needs of people in our society, it seems unfortunate, if not misguided to build a testing program on the content of those programs. To build a structure of testing related to a status quo entrenches an establishment. Whether the materials of instruction should be built on the tests (the perennial criticism of what has emerged from the New York Regents' exams,) or whether the tests should be built on the instructional materials, is only the most visible part of the issue. What is basic is that *both* the tests and the materials should relate to the most important needed learning.

Tests must be based on something; and in the absence of widely agreed-upon specifications of objectives, the most widely used textbooks constitute, by default, a basis for test design. Perhaps there is no practical way to get broad-scale agreement on objectives, but the profession should be alert to the fact that testing formalizes implicit objectives, making them more visible, more controversial, and hopefully, more apt to be refined. Current emphasis on testing programs offers both challenge and hope. Confidence and hope for constructive impact of a testing program rests largely on the assumption that tests are testing the things that are truly important. Thus, confidence may now be lacking to the extent that there are some profound misgivings about the adequacy of the current crop of textbooks and instructional experiences within the school.

It is indeed a sad commentary on society that the most basic human issues seem receive so little educational attention. Perhaps as we see our efforts in the light of more careful assessment, this will finally come home to us. After three years of near-panic over the tragedies brought on by our denial of basic human rights (and having achieved no profound closure,) we now rush into a “new” anxiety: ecological disruption. How characteristic! What even *Time* magazine has tagged “Nixon’s New Issue” has been called to our attention for many years by conscientious, concerned men and women; but now that it is a “technological” problem—now that the engineers and the technologists have decided their services are needed—it is suddenly dubbed an *official* problem. Perhaps it’s a safer problem because it seems to deal with *things* more than with people and their beliefs. But does it? Neither the problem of human carelessness nor the integrity of the environment is apt to be solved technologically. Further, these problems are not apt to be solved by schooling alone. The forces in society that teach us to place highest value on convenience and fun are far too formidable. Yet we

¹ Presentation: The Michigan School Testing Conference Rackham Building, The University of Michigan, Ann Arbor, Michigan, February 18, 1970. Sponsored by The University of Michigan Bureau of School Services with the co-operation of The University Extension Service.

hear of people and agencies at the state and national level preparing to relieve the national conscience by requiring the schools to add units and courses on environmental quality. What is there about our outlook, our perception, and our consciousness that is so cognitively oriented, so fixed on facts that we leap to the proposition that any evil can be relieved by enough of the right information? Do the Dutch keep their flat, and potentially grim, grimy land colorful by teaching their children courses in flower planting? Do the Swiss keep their crowded rolling plains an inhabited showplace of the world by teaching units on public neatness? Indeed not; for centuries the traditions have been nurtured by public awareness and private conscience. Schools play a part, indeed, but they work hand-in-hand with larger expressions of a commonly held value system. Schools, especially schools obsessed with a cognitive emphasis, are a poor substitute for universal concern and universal responsibility about the rights and needs of humanity.

If this seems a digression, I beg your indulgence to see what the next few years will hold. If the large-scale movements toward wider use of testing go one way, we win; if they lead another way, we lose. The most profound challenges in our society are not cognitive. If testing makes us more aware of this, well and good. If testing makes us more preoccupied with cognitive learning, schooling will become less and less relevant.

Tests, and especially testing programs that allow for comparisons among teachers or schools, exert considerable influence on the curriculum. Under certain circumstances this influence can be advantageous, but not always. Those who determine the tests to be used in a given instructional system hold a vital control over that system, whether or not they recognize it. The relative emphasis between effective and cognitive learning illustrates this power rather well. Testing, because of human tendencies to opt for the easier, can increase the emphasis on cognitive learning. As educated people—better to say schooled people—ourselves, we are more aware of the functions of knowing (recall, recognition, association) than the functions of feeling and being. The management of cognitive information (facts, figures, dates, sequences, procedures, and techniques) commands our primary attention as teachers. But all the while, our students seem to be concentrating on liking, disliking, affiliating, disaffiliating, and other matters more of *being* than of knowing. It's as if teachers and students exist in separate worlds.

Recently, a high school teacher described affective objectives as “spooky”—that he felt more at home spelling out cognitive objectives. This teacher is not unique and he is right in his presumption that cognitive matters are easier to specify and to test. Our system is more strongly affected by what is easiest, than by what is best. As teachers and curriculum workers, we need help from test designers; we can't afford to let over-emphasis on cognitive learning further isolate us from the affective world of our students. We must work together to find better ways to specify objectives and develop tests for the affective changes our efforts produce in learners. For example, as the ecology issue emerges, let's get hold of it both in the affective and in the cognitive domains. What a precedent we could create if we insisted that every unit of studies in environmental quality should have clearly defined affective objectives as well; and that every test in this field should test for affective learning!

I commend to your attention the thoughtful and stimulating April 1969 issue of *Theory Into Practice*. Jack Frymier pulled together many pieces relating to the whole problem of teaching the values that play a greater role than does “knowledge” in determining human behavior. Thorough behaviorists owe it to themselves not to be ignorant of the part played by values and beliefs.

Another concern is that testing can tend to produce blind spots in evaluation. This contention may be rather obscure, but perhaps we can see it this way: Test items are ordinarily representations of instructional objectives, explicitly and implicitly held. Instructional objectives are invariably expressed in terms of what we want to happen. Thus, tests are characteristically limited to assessing the degree of presence of some desired behavior. Since some changes induced by

schooling are negative, ordinary tests can completely fail to pick them up. Failure to assess undesirable learning is an important oversight. Here, too, we need creative help from test designers.

Lest you suspect that I have come here not to praise testing but to bury it, let me quickly assert that from an instructional systems viewpoint, testing is an absolute necessity. It is, or can be, a stimulant for change and improvement.

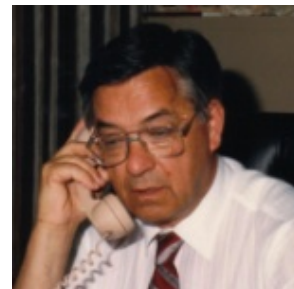
Most of you are personnel specialists or other than classroom teachers. Much of what I have said here today speaks primarily of the classroom aspects of the curriculum. Regardless of how you define “curriculum,” the action is where the teacher is. I believe that those of you who do understand testing—its potentialities and its limitations—have a job to do. Someone needs to get through to teachers about what testing can mean in terms of improvement of the curriculum. The issues of testing, state and national assessment, and so forth, have been unfortunately polarized. You are “all out” for it or “all out” against it. Perhaps the only way we can get together is to focus on better educational experiences for boys and girls—and then to discover the vital roles that testing plays in stimulating, identifying, and quantifying improvement.

In an instructional system, testing is essential because it produces data to serve as a basis for decision-making. Data is essential if decision-making is to be anything but intuitive. Highly insightful teachers, especially in earlier and less complex eras, made intuitive educational planning a fine art. For various reasons, the scene has changed. The quality of educational planning is somewhat out of joint. A restive, over-taxed, and anxious public seeks scapegoats. Yet in a relatively informed society, liquidating a scapegoat that may be needed again isn’t reasonable; instead, he is made to “toe the mark,” preferably for “his own good.” So now the name of the game is *accountability*. Although we may question the motives behind much of the pressure for accountability, there is something here that can be turned to professional advancement.

As it stands, teachers are apparently more able to “sense” things than to *see* things. A recent study in the Texas Research and Development Center for Teacher Education doesn’t do much for the hope that teachers are being accurate in their sensing of what children are capable of being and doing.² The NCME Measurement in Education Series, if the quality can continue to match that of the first issue,³ should constitute a helpful resource. Thus, I am especially grateful for materials that help teachers become more able to appreciate and wisely use tests.

Accountability, in professional terms, is the evaluation of system outcomes in light of objectives and costs. There is reason to believe that increased public confidence in the schools would follow from sincere efforts to assure that what is supposed to happen in schools is indeed happening. People are entranced by the idea that learning can be made subject to the terms of a contract. Illustrations are appearing frequently in recent months. Don Davies, Bureau Chief of the Educational Personnel Development Act, moved “accountability” into the spotlight with his pronouncement last fall: “We are moving in a direction we have been contemplating for a long time—shifting primary learning responsibility from the student to the school. It also means a lot of people are going to be shaken up.”

Ralph Becker of the Bureau of Elementary and Secondary Education describes how this shift is being implemented: “Agencies [must] accomplish what they say they are going to accomplish in their proposals within a given time. If these specifications are not met, a penalty may ensue.”



² Thomas Good et al.. Teacher Assessment of Pupil Potential Report (Austin, Texas: Research and Development Center for Teacher Education, Texas University, 1969)

³ Robert Thorndike. Helping Teachers Use Tests. National Council on Measurement in Education, Office of Evaluation Services. (East Lansing, Michigan: Michigan State University, 1969)

The director of a dropout prevention program in St. Louis has pledged to reduce the dropout rate by 10% in the schools where his OE-sponsored program operates.⁴ Lee Wickline of the Division of Plans and Supplementary Centers talks about one form of penalty: “If a project director proposes to continue the same processes the second year that proved ineffective the first, the likelihood of his receiving funds would certainly be less than if he had succeeded or if he had proposed to change his methods.”

The “Yardstick Project” in Ohio reflects this sort of thinking but without the federal clout. Their watchwords are “growth gauge” and “effectiveness of decisions.”

Whether industry’s newfound role in education is the cause or effect of this surge toward accountability, the historians will have to judge; but it’s in there somewhere. By now, Dorsett Educational Systems’ Texarkana Project has put both Dorsett and Texarkana on the educator’s map. Their innovation is the “guaranteed performance” contract. The achievement gains of children will determine the amount of payment Dorsett collects. Preliminary findings, according to Dorsett’s own impartial surveys, strongly suggest that the old extrinsic rewards argument has been settled for once and for all.

San Diego has called in McGraw-Hill’s Educational Development Laboratories on a similar basis. The *Washington Post* called it a “hard-headed approach to upgrading the nation’s schools.”

To gain perspective, consider what evaluations have been showing. For example, the National Advisory Committee on Dyslexia and Related Reading Disorders, upon completion of a year of study, reports a sort of “directionless shambles” posing as reading instruction across the nation.⁵ The Commission on Instructional Technology reports waste and failure in reference to one after another of the efforts to improve teaching through technology.⁶ These harsh indictments are neither scarce nor, apparently, unfounded.

Much of the current enthusiasm for making important ideas work can energize professional improvement. Surely a knowledgeable use of testing will play a key role in turning the concept of accountability from a threat to an impetus.

Consider the following five suggestions about the curricular function of testing:

1. Examine and reflect on test items:
 - What do the items suggest about what is really being valued by the educational establishment?
 - What relationship exists between what is being taught and what is being tested?
2. Consider any dissonance among objectives, teaching, and testing to be a mandate for change.
3. Continually challenge the validity of the educational objectives in the system; keep them specific, valid, and relevant to the needs of society.
4. Strive for balance among the learning outcomes to be tested; avoid undue emphasis on those things that are merely easy to test.
5. Value evidence about what is being accomplished by teaching.

To recapitulate:

1. Testing makes teaching objectives more formal, more visible, more controversial, and, with leadership, more apt to be refined.

⁴ Startling Reading Gains Hinted in Texarkana. *Education U.S.A.* (February 2, 1970)

⁵ Reading Disorders in the United States. United States Office of Education.

⁶ To Improve Learning. United States Office of Education.

2. Testing provides data on the effects of an educational experience or system.
3. Testing provides data on the learner, allowing an attunement between his or her learning (in terms of capability, readiness, needs, and progress) and the instructional experiences provided by the system.
4. Testing provides a control component—an informal or formal force for conformity of individual instructional units to the goals of the system.

Two specific cautions have been suggested:

1. Testing can increase the emphasis on cognitive learning. (While this is not inherently wrong, it certainly can get out of hand.)
2. Testing can tend to produce blind spots in evaluation—especially with reference to failing to identify the unwanted learning that occurs.

A more accountable and responsive establishment could result if, in addition to the needed data on learners, their capabilities, readiness, needs, and progress, there would also be an emphasis on testing for system evaluation: Do the instructional procedures and materials teach? What? How much? To whom? If testing is to have a positive impact on the curriculum, teachers and administrators must be carefully tutored to accept less defensively and to *utilize* feedback about the effects of their efforts. If teachers could get this frame of mind, it would make a substantial difference in the quality of schooling. We would see more attention to refinement of teaching, more attention to definition and validation of objectives, more majoring on the important things, less confusion and haphazard slippage.

But we have a long way to go.

References

- Education U.S.A.* "Startling Reading Gains Hinted in 'Texarkana,'" Washington, D.C. (February 2, 1970.)
- Good, T.L., D.L. Williams, R.F. Peck, and L.M. Schmidt. Teacher Assessment of Pupil Potential. The Research and Development Center for Teacher Education, The University of Texas at Austin, Report Series No. 33 (September 1969)
- Theory Into Practice, Teaching the Young to Love* (entire issue) Vol. VIII, No. 2 (April 1969)
- Thorndike, R.L. Helping Teachers Use Tests. *National Council on Measurement in Education*, Vol. 1, No. 1 (October 1969)
- U.S. Department of Health, Education and Welfare, Office of Education, New Programs Face Stricter Evaluation. *American Education*. Vol. 6, No.1 (January-February 1970) p. 36
- U.S. Office of Education. *To Improve Learning*. Washington, D.C.: Government Printing Office (1970)
- U.S. Office of Education. *Reading Disorders in the United States*. Washington, D.C.: Government Printing Office (1970)
- Yardstick Project, Alcazar Hotel, 2450 Derbyshire Road, Cleveland, Ohio 44106