

- Can we depend upon objective tests and other methods of valuating educational achievement and knowledge?

PROSPECTS FOR EVALUATION OF LEARNING

What of the prospects for educational evaluation? Will the present practices be reversed? Will the present problems be resolved? The long history of education suggests that enduring changes are more likely to evolve slowly than to explode suddenly. But changes do come.

One of the current and anticipated changes has to do with the increased emphasis on education and its evaluation. Since World War II, the rush of students to college, in greater numbers than most good colleges could accommodate effectively, has led to enormous expansion of admissions testing programs. The flow of dollars to aid students who are able but not affluent has led to the development of scholarship testing programs. The needs these testing programs have served will continue, and no better alter-

native seems likely to develop. But we ought to hope and expect that tests will improve and their results will be used with increasing wisdom.

New concern for quality in education and for equality of educational opportunity, with resultant increases in government expenditures and involvement, have led to recognition of the need for reliable assessment of the results of our educational efforts. The national assessment is one attempt to meet this need, and some states have enacted law relating to mandatory testing and reporting of test results in the public schools.

The growth of the wide-scale programs for testing educational achievement and for college admission and scholarship testing has led to another major change: the development of high

speed, high capacity, highly automatic machines for scoring objective tests. It is reasonable to suppose that the years ahead will see rapid growth in the utilization of these diverse and versatile devices.

Despite the current popularity of objective tests and mass testing programs, however, there are those who believe that all is not well with the evaluation of learning today. Among the concerns expressed are these:

1. *That the tests currently used to evaluate learning are inadequate to the task, measuring only imperfectly the less important educational outcomes.*

One's opinion on this matter is, of course, likely to depend largely on whether he agrees with what most schools spend most of their time trying to do, that is, to help students gain command of useful verbal information. The subject matter of most studies — history, literature, science, geography, even mathematics (if its symbols are regarded as essentially verbal symbols) — is verbal information. If verbal

information is extracted from formal education, there is very little if anything left.

But many educators are unwilling to admit that their aims are so prosaic, preferring to claim objective that are more spiritual than material, and hence largely immeasurable. Verbal knowledge is certainly not all that matters where man is concerned, and the school cannot afford to ignore muscular skills, or attitudes, or values, or character, or overt behavior. But neither can the school afford to give any of these things priority over command of knowledge in specifying its mission. And if it should choose to give other things priority, it will almost certainly find that cultivation of command of useful knowledge is the best, if not the only, means it can use to attain the ends it seeks.

If this is true about man and the process of educating him, tests can do much of the job of evaluating learning. Many tests in current use are inadequate, it is true. But their faults lie less in the direction they point

than in the distance they travel.

OBJECTIVE TESTS

2. *That objective tests are spuriously attractive because of the ease with which they can be scored en masse, but are seriously deficient as tools for the evaluation of learning because of their inherent ambiguities, their tendency to emphasize superficial factual information, and their reward of successful guessing.*

The supposed deficiencies of objective tests, however, are not inherent in the form. Objective test scores are typically more reliable than essay test scores, both because each student's performance is judged against the same standard, and because of extensiveness in sampling various aspects of achievement. It is true that objective test questions appear to be trivial more often than do essay test questions, but this is a matter of numbers. If a test can include only a few questions, as an essay test ordinarily does, the tendency is to make each one general and comprehensive.

Objective test questions also tend to be more "factual," but it is important to remember that a fact in this sense is a verifiable truth, which need not be trivial. If a subject is not loaded with important factual truths, the value of studying it would seem open to serious question.

It is true that answers to objective test items apparently could be learned by rote; without real understanding, but this seldom happens. For one thing, it is always possible to pose questions that the examinee has never encountered before, and thus require answers he could not have learned by rote. For another, rote learning is a difficult, ineffective, and unsatisfying method of learning most things that students study.

That test questions, either in objective or essay form, are sometimes ambiguous is also beyond dispute. But with reasonable skill and care in test construction, this can be reduced to the point where it no longer interferes seriously with the evaluation of learning.

Like ambiguity, guessing is not a genuine menace in the use of objective tests. Well-motivated students do very little blind guessing on tests that are appropriate for them. The correctness of their informed guesses is related substantially to the amount of relevant information they command. Thus their "guesses" provide valid indications of achievement. A student who does a great deal of blind guessing is likely to get a very low score on a good test. Finally, both ambiguity and guessing would result in inconsistent results from repeated measurements, and so if a test constructor succeeds in building a test that yields reliable scores, it is safe to conclude that defects related to ambiguity and guessing are not serious on that test.

Thus despite the criticisms of objective tests, it seems likely that their popularity will continue to grow.

3. *That wide-scale testing programs and the use of standardized tests place teachers in curricular strait-jackets, preventing them from meeting local needs or mak-*

ing use of unique local opportunities, suppressing their creative ideas and their individualities as teachers, and rewarding routine, mechanical teaching.

It is true that if students and teachers know in advance the general nature of questions to be asked and content to be covered in a test used to evaluate learning, they will direct their study and teaching toward these kinds of capability. But if the tests are good tests, with appropriate curricular coverage and emphasis, and if they are not the sole basis for evaluation, they are likely to do much more good than harm. After all, the test-makers, in most cases, are themselves master teachers, and the tests they build aim to follow rather than to lead curricular innovation. The teachers most likely to make the review of old tests a major part of their instructional program, as if they had been placed in a curricular straitjacket, are those who are least secure in their positions because they are least competent.

External tests have been influencing what is taught in particular classrooms for nearly 40 years; yet is it not true, in view of the increasing mobility of our people, that a greater degree of uniformity among classrooms than we have today could well be tolerated?

4. *That testing places students under undue pressure and exposes them to unnecessary experiences of failure, diminishing their self-confidence and destroying the joy of learning.*

It is not the measure of achievement but the aspiration to achievement that places students under pressure. Test scores simply report levels of achievement; if the reports are disappointing, the blame may rest on ineffective learning or teaching, or on unrealistic expectations.

The suggestion that the way to deal with excess pressure is to stop paying so much attention to achievement makes very little educational sense. Instead we need to pay more attention to the setting of realistic goals, and to the recognition

of individual differences in interests, abilities, and avenues for self-fulfillment.

5. *That testing, particularly intelligence and aptitude testing, leads to the labeling of pupils as bright or dull, in both cases adversely affecting their expectations, their efforts, and their self-concepts; denying and thus tending to destroy the almost infinite potential for development inherent in every human being.*

Although the items in most intelligence and aptitude tests are clearly measures of developed ability, too many educators have been willing to believe that they provided direct and dependable measures of innate capacity for learning. On too many occasions, a child's low IQ score has been used to explain his failure to learn instead of being used to help him to learn.

But these tests have sometimes been interpreted properly and used constructively. It is hard to beat a good intelligence test as a convenient measure of a young child's general educational development. Since

all learning builds on prior learning, effective teaching requires information on each child's level of educational development.

More schools may join those which have abandoned intelligence testing because of abuses and because of local pressures, but it is not likely that intelligence testing will disappear. We can hope and expect, however, that intelligence and aptitude tests will be interpreted more

realistically and used more constructively.

In all, to teach without testing is unthinkable. Teachers are likely to do more testing in the future, and to do it better as they become more skilled in the techniques of their craft. Above all, they are likely to use the results of testing more wisely and more constructively. — *By Robert L. Ebel, from The Education Digest, March, 1969.*

AWARENESS OF LIMITS

As we advance in life, we learn the limits of our abilities. — *James Anthony Fraude*