Curriculum Development Evaluation: The Need To Look Beyond Behavioral Objectives

DONALD HOGBEN

Donald Hogben is professor of education at Flinders University of South Australia.

The characteristics of appropriately worded behavioral objectives and the advantages for curriculum design and implementation of a clear specification of objectives in advance of any teaching or testing have been articulated by a number of people, for example, Mager,[1] Popham,[2] and Sullivan.[3] Essentially, a behavioral objective is a statement or description of intent. It is not, however, a statement of what a teacher intends to do, but rather, a statement of what the teacher intends that the student will be able to do or produce at the conclusion of some period of instruction.

A properly stated behavioral objective must describe without ambiguity the nature of learner behavior or product to be measured.[4]

Two major advantages are claimed for behavioral objectives. First, they provide clear end points toward which all can strive; and second, because they focus on expected terminal performance of students (what students are expected to be able to do), they suggest methods of assessing the extent to which objectives have been realized. The apparent logic of such an approach is obvious to all; to argue against behavioral objectives would seem to be to argue for ambiguity, if not irrationality. Nevertheless, a number of people have drawn attention to some of the difficulties and possible hazards of the approach, for example, Atkin,[5] Eisner,[6] and this author.[7] It is not my intention here to go over old ground; however, I do wish to draw attention to some very serious dangers in evaluating programs from the simple instructional model implied in the behavioral approach.

In this article, I intend to deal with the danger of assuming that, because we devise measuring instruments (most often paper-and-pencil tests) to assess the degree to which stated objectives have been achieved, we are in any way getting a reliable and valid picture of the outcomes of an instructional sequence. At best we are assessing, with varying degrees of accuracy, how well students can do those things identified by us at the outset as things that we wanted them to be able to do. The very process of goal identification and statement is too often constrained by a narrow perception of achievement measurement.

The behavioral objectives approach, by definition, calls for a high degree of isomorphism between objectives and test items or tasks.[8] If we cannot state without ambiguity what the student shall be able to produce, then we are not stating an objective properly in behavioral terms. And, because one must avoid inferences about mental events, the "production" must be overt, observable. The following objective test item, provided by Sullivan, nicely illustrates these features.

Objective: Given a list of instructional objectives, the learner will identify those that are stated in terms of learner behavior.

Test Item: Mark an X by the number of each behavioral objective in the following list.[9]

Let us suppose for a moment that we are able to state all our intents in the form of highly specific behavioral objectives of this sort before any instruction commences. If we can, we will have before us a clearly defined universe of intended student behaviors. The degree to which a student can demonstrate these behaviors can be taken as a measure of his success. So far so good. But can success by all students (or some agreed-upon proportion of them) be taken as a reasonable measure of the success of the course or program? Well, yes it can, but only if all the intents were stated at the outset (and we agreed that they were worthwhile intents for all students to pursue); all the intents were measured (or an accurate and unbiased sample of them); and there occurred no

unplanned for, undesirable outcomes. If we are evaluating a course or program, we might settle for the achievement of stated objectives if, and only if, there is congruence between worthwhile intents and actual outcomes. But how well is this requirement satisfied in practice? How well might it be satisfied?

Anyone who has made a really serious attempt to state course goals in behavioral terms will immediately recognize this to be an enormously time consuming and difficult operation. We have so many goals; we hope to achieve so much. We have intentions of which we are only vaguely aware. We have aspirations for our students that we hope can be fulfilled, but we wonder about our chances. We have goals that are of great importance; we have others that are of only marginal value. Where do we start?

Let us start by first writing out all our goals as quite general statements. We could then take each goal in turn and break it down into a number of more specific statements, each worded in terms of measurable student performance.10 Having attempted this, we very quickly come to realize that, for a course of any substance, we are going to finish up with a huge list of statements, perhaps hundreds. Another thing strikes us as well. Translating some of our goals into behavioral objectives is really not all that difficult, but there appear to be so many that just seem to defy such translation. Also, we are more than a little concerned to note that the goals that give us the most bother are very frequently those to which we had initially awarded high priority. We may well find that a number of our high-priority goals simply cannot be translated into behavioral objectives. Does this mean that these goals are unrealistic; should they be deleted from our list? Perhaps. Are they, however, worthwhile intents that should be retained even though they cannot be operationalized? Very likely. Unfortunately, what most often appears to happen in practice at this point is that these goals are classified "unrealistic" and get omitted, and the final list becomes heavy with cognitive objectives that emphasize the memorization and recall of information or the mastery of relatively simple skills. These are the easiest objectives to identify, the easiest to measure using standard procedures, and hence the easiest to state in behavioral form. If the final list of course objectives omits valuable, high-priority goals because they defy translation into behavioral objectives, then the achievement by students of the listed behavioral objectives cannot be taken to mean that they have satisfied the goals of the course.

ACHIEVEMENT TESTS AND EDUCATIONAL GOALS

Consider the following statement concerning goal definition by Henry Dyer of the Educational Testing Service:

. . . in the last analysis, an educational goal is defined only in terms of the agreed-upon procedures and instruments by which its attainment is to be measured. It is to say that the development of educational goals is practically identical with the process by which we develop educational tests. It is to imply what in some quarters might be regarded as the ultimate in educational heresy: teaching should be pointed very specifically at the tests the students will take as measures of output; otherwise, neither the students nor the teachers are ever likely to discover where they are going or whether they are getting anywhere at all.11

Such a statement appears to be wholly consistent with the behavioral objectives approach. Tests are clearly seen as the devices by which objectives are defined and by which terminal behavior is measured. It seems that educators have become so locked to the paper-and-pencil test (perhaps more specifically to the multiple-choice item) that educational achievement and performance on such tests has become almost synonymous. The grip of the "objective" test is so strong, the technology of testing so well developed, that it appears to have set the rules and defined the boundaries within which educators have to operate. People have become so familiar with such instruments, feel so comfortable with them, that minds often seem closed to other possibilities. Translating a general goal into specific behavioral objectives (what the student should be able to do at the end of some sequence of instruction) has come to mean, for many people, translating goals

into objective test items. The translation task is often so difficult, however, that many valuable goals get lost in the process. Dyer himself is not unaware of the difficulties and hazards, as he follows the passage quoted above with the following:

A great problem—probably the greatest problem—in the development of meaningful goals is that of making sure that the tangible tests that come out in the process bear a determinable relationship to all the vague individual and collective concerns that go into it.12

If the behavioral-objectives approach is followed with fidelity, as things stand at the moment, there will be a high degree of relationship between performance measures and stated objectives. But it is almost certain that the sum of the stated behavioral objectives will be less than the total of the original goals that generated them. Using this model means, then, that, although we can assess with some accuracy the achievement of the stated objectives, it is highly unlikely that we will obtain any reasonably valid assessment of student achievement based on original intent unless, of course, our original goals were highly performance oriented. Further, if assessment is limited to stated objectives, we have nothing approaching a full evaluation of the course or program. An evaluation requires not only the assessment of all the intended goals, but also some attempt at the assessment of actual program outcomes.

If we could translate all our goals into highly specific behavioral objectives, then, because of the isomorphism between objectives and test items or tasks, we would essentially be constructing our measuring instruments as we write out our objectives.13 The validity of the content of our final tests should be perfect because our items are our objectives; our universe of objectives and our universe of test items are congruent. In practice, however, the sheer number of statements and hence test items required to achieve this ideal is just too great. So even if we could operationalize all our goals in this way, it is doubtful whether it would be a practical undertaking. However, what I want to suggest here is that, because we tend to think of achievement in terms of performance on paper-and-pencil tests, and because of the not inconsiderable imperfections and shortcomings of these instruments, a danger at this time is in fact thinking in terms of this ideal.

The psychometric test is so much a part of the makeup of the educator that in the process of goal identification and statement he is very likely to shape his goals to fit the tests. Most times he may not be conscious of doing this, but I am sure he does it nevertheless. Now, if our measurement instruments were perfect (or very nearly so), this might be a sensible thing to do; this, I think, is the point that Dyer is making. However, at the present time our instruments are such that to manipulate our goals to fit them would be to grossly distort the full intent of most programs. And to assume that performance on such tests can be taken as any reasonable assessment of course outcomes, and thus serve as an evaluation of program worth, is at best naive. In the words of Hastings:

. . . we as a group, having adopted the techniques of psychometrics and experimental design, tend to be more concerned with altering the problem so that it can be tackled by these techniques than we are with adopting and adapting techniques for attack on the complex problems. An example of this is our general insistence on approaching the problems of evaluating educational endeavour by developing a priori statements of behavioral outcomes, applying our psychometric techniques, and then acting as though we had looked at the whole bit.14

INSTRUCTIONAL IMPLICATIONS

Assume for the moment that we do write out all our behavioral objectives. (We have plenty of paper and lots of time.) As I have pointed out above, the isomorphism between objectives and test items guarantees the validity of the tests we develop to assess the extent to which students have achieved our objectives.15

Let us assume also that the reliability of our measurement is going to be high. Before the course commences, then, we have clearly before us what we want our students to be able to do at the end:

our goals are clear and fixed; our tests are all but constructed. What is the focus of learning within such a framework? Several things are clear. First, at the end of the period of instruction, and after we have scored our tests, we will know whether or not students can do those things we had planned that they should be able to do. But we will not be able to say whether they can do anything beyond that which we had planned for them, whether they learned anything beyond that which we had stipulated at the outset. Second, because everything has been set out in advance, the whole emphasis must be on teaching and learning the known. Even what is to be "discovered" must be preset: student imagination, curiosity, inventiveness must be carefully guided toward the discovery of what was set out in detail in advance. It is a closed instructional and learning model. Atkin has drawn attention to a number of the shortcomings in the behavioral-objectives approach associated with curriculum planning and implementation.16 He points out that teachers need considerable flexibility, as the most appropriate time to teach something is often not that planned for it in advance. Stake has discussed the same general idea, and he notes that:

There are many advantages to external programming, e.g., writing lessons in advance as the programmed instruction people do or as the well-organized lecturer does, but these advantages should be weighed against the advantages of assigning control to teachers who are sensitive to conditions optimally suited for the pursuit of elusive long-range goals.17

Third, although the model can easily accommodate to individual student differences in terms of learning speeds and preferred learning modes, it does not cater to individual differences or preferences with respect to what might be learned. This is particularly true when students as well as teachers must operate within a framework of objectives stated unambiguously in behavioral terms. Surely teachers hope to develop in at least some of their students an interest which manifests itself in learning beyond that demanded for "mastery." As I see it, there is no room in the behavioral model for teaching that seeks to encourage and reward independence of thought, originality, initiative, and imagination. Vague terms? Certainly they are, but this does not negate their value.

At several points in this paper I have made the distinction between actual course outcomes and the achievement of intent. This distinction is vital, but blurred or overlooked so often that it deserves some further elaboration. It is more than a little likely that the behavioral objectives model, because of its apparent precision and certainty, may lead us into this trap more often than less structured and more open approaches to instruction and student assessment. Moreover, I have argued that the translation of general goals into specific behavioral objectives is likely to result in a de-emphasis on or omission of a number of worthwhile intents at the outset. If this does occur, then even the accurate assessment of objectives falls short not only of outcome assessment but also of full assessment of original intent,

Atkin and Dyer, among others, have drawn attention to the necessity for evaluating the side effects of a program, pointing out that these may be of more significance than the intended effects.18

An up-to-date math teacher may be trying to teach set theory to fourth graders and may be doing a good job at it, but one wants to know whether he is also teaching some of the youngsters to despise mathematics.19

While I heartily agree with such sentiments, I would also argue that such side effects, or rather the reverse of them, should be elevated to the status of intended main effects. That the child should not despise mathematics (or literature, or physics, or whatever) is just as important to my mind as the achievement of a good many of the things that one typically encounters in lists of behavioral objectives. The worth of a program that discourages originality, imagination, independence of thought, active and aggressive questioning, etc., must be suspect. Today, perhaps more than ever before, we need to help students acquire (or retain) the intellectual tools and emotional capability to deal with the unknown and the changing. I do not believe that at the present time insistence on behavioral objectives is the answer. Nor do I believe that the statement of goals in general terms will necessarily lead to more creative, innovative, and appropriate teaching. What I do believe we

need at this time is a lot more focus on the actual outcome of education and less attention on the explicit formulation of intent.

The total impact of many programs or courses has been assessed often simply as the difference between pre- and post-test scores on some achievement test. No statistically significant difference is taken to mean no change (and presumably course failure); a gain of one or two grade levels is hailed as success. Such change, or absence of it, is only part of the total picture. There is much besides that needs to be assessed and evaluated. The total outcome of any course or program is the sum of many changes; and this sum includes not only planned-for acretions, but also deletions and retentions (intended and unintended) as well as unplanned for gains.

The final equation is indeed a complex one, but it is one we must come to terms with. There are, for example, many characteristics of the average five- or six-year-old entering school that we would probably not want to change but retain: happiness, curiosity, honesty, optimism, excitement. How often do we look for the retention of these things in our evaluation of curricula? Many techniques remain to be developed for the assessment of outcome. There are many outcomes that still remain to be identified. It is here that our efforts should now be concentrated.

1 R. F. Mager. Preparing Instructional Objectives. Palo Alto, Calif.: Fearon Publishers, 1962.
2 W. J. Popham, "Objectives and Instruction," American Educational Research Association Monograph Series on Curriculum Evaluation, Vol. 3. Chicago: Rand McNally, 1969, pp. 32-64.
3 H. J. Sullivan, "Objectives, Evaluation, and Improved Learner Achievement," Ibid., pp. 65-99.
4 Popham, op. cit.r p. 37.
5 J. M. Atkin, "Some Evaluation Problems in a Course Content Improvement Project," Journal of Research in Science Teaching, Vol. 1, 1963, pp. 129-132; and J. M. Atkin, "Behavioral Objectives in Curriculum Design: A Cautionary Note," The Science Teacher, Vol. 35,1968, pp. 27-30.
6 E. W. Eisner, "Educational Objectives: Help or Hindrance?' The School Review, Vol. 76, 1967, pp. 250-260; and E. W. Eisner, "Instructional and Expressive Objectives: Their Formulation and Use in Curriculum," American Educational Research Association Monograph Series on Curriculum Evaluation, op. cit., pp. 1-18.
7 D. Hogben, "The Behavioral Objectives Approach: Some Problems and Some Dangers," Journal of Curriculum Studies, May 1972.
8 Eisner, "Instructional and Expressive Objectives: Their Formulation and Use in Curriculum," op. cit.
9 Sullivan, op. cit., p. 75.
10 See D. R. Krathwohl, "Stating Objectives Appropriately for Program, for Curriculum, and for Instructional Materials Development," Journal of Teacher Education, Vol. 16, 1965, pp. 83-92; and P. A. Taylor and T. O. Maguire, "A Theoretical Evaluation Model," Manitoba Journal of Educational Research, 1966, pp. 12-17.
11 H. S. Dyer, "The Discovery and Development of Educational Goals," in J. C. Stanley, Chairman. Proceedings of the 1966 Invitational Conference on Testing Problems. Princeton, New Jersey: Educational Testing Service, 1967, p. 22.
12 Ibid.
13 See M. Scriven, "The Methodology of Evaluation," American Educational Research Association Monograph Series on Curriculum Evaluation, Vol. 1. Chicago: Rand McNally, 1967, pp. 39-83.
14 J. T. Hastings, "The Kith and Kin of Educational Measurers," Journal of Educational Measure ment, Vol. 9, 1969, pp. 127-130.
15 Teaching the actual items would, of course, destroy validity except where the objective calls for the recall of specific factual information or the performance of some fairly simple skill.
16 Atkin, "Some Evaluation Problems in a Course Content Improvement Project," op. cit.; Atkin, "Behavioral Objectives in Curriculum Design: A Cautionary Note," op. cit.
17 R. E. Stake, "Language, Rationality, and Assessment," in W. H. Beatty, ed. Improving Educational Assessment and an Inventory of Measures of Affecting Behavior. Washington, D.C.: Association for Supervision and Curriculum Development, National Education Association, 1969, p. 33.

18 Atkin, "Behavioral Objectives in Curriculum Design: A Cautionary Note," op. cit.; J. M. Atkin, "On Looking Gift Horses in the Mouth: The Federal Government and the Schools," The Educational Forum, Vol. 34,1969, pp. 9-20; Dyer, op. cit.
19 Dyer, Ibid., p. 20.