BARBARA BIBER

Problems of Values and Measures in Evaluation of Teaching

This author illustrates, from practice and research, problems encountered and a tentative pattern of solutions adopted in a program of evaluation of teaching.

THE PROBLEMS of evaluation of teaching rest upon a question of viewpoint. That this is true at the point of defining goals is broadly recognized. That it is also true at the stage of selecting and developing methods of evaluation is more likely to be only vaguely realized or missed altogether. Perhaps it is safe to say that we can look for significant insights to those situations where there is conscious awareness of the need for a consistent framework of values, in this instance, a clearly defined concept of the teacher's role, to be applied at all stages of evaluation. In the practical world, the problem is obvious: whether a training school's opinion of a student is borne out after he is teaching is often, in large measure, dependent on which superintendent of schools is judging the teacher. In systematic research, how correct a series of predictions proves to be is, to an important degree, a function of whether the criterion measures are conceptually coordinate with and relevant to the predictive variables.

In the brief space of this paper, I should like to illustrate, from practice and research, the problems encountered and the pattern of solutions adopted (tentatively to be sure) in a program of evaluation which proceeds from a viewpoint and attempts to operate consistently within it.[1]

The impact of John Dewey's social philosophy for a democracy moved the goals of education toward emphasis on active participation in the learning process, independent and responsible thinking and functioning and a creative-constructive relation between the individual and the group. Later, the psychodynamic interpretation of human behavior contributed other dimensions to these goals: the interdependence of the processes of emotion and learning, the importance of the teacher-child relationship as proving ground and prototype of interpersonal relationship experience in general, the adaptation of content and methods of curriculum to the changing needs, drives and conflicts associated with the succession of stages of maturity.

Out of the amalgam of these major influences, two basic goals can be formulated and stated very briefly. First, it is the goal of education to help the child

[1] The program to be drawn upon for illustration is in operation at the Bank Street College of Education in New York City.

***BARBARA BIBER** is director of research, Bank Street College of Education, New York, New York.*

to become skilled, knowing, perceptive and effective in his world[2]; to master confusion; to reorder his experience through his own invention and creation; to communicate in the modes of his culture and sustain the idiom of his own individuality; to extend his "interest" world to far places and times; to act, to organize, to accomplish, to reason, to reflect. As important as the goal itself is the additional responsibility to deal with the child's learning, achievement and mastery in such a way that they constitute the media for positive ego-growth and thus contribute to the child's growing feeling as a knowing, doing, confident self in relation to a knowable, manageable world.

Secondly, it is the goal of education to contribute to the child's increasing maturity and satisfaction in his modes of relating to people; to make it possible for him to find in his teachers adults who can connect with him meaningfully as a person, respect his distinctness, be attuned to his feelings as well as to his capacities and be aware of the importance of his private world to him. In such a relation with his teachers he will experience again, significantly in a new context and with different emotional overtones, what he has already lived through with his parents: the taste and the boundary of freedom; the comfort and irritation of being controlled; the safety and threat of being known. By the way these relations are mediated for the child and by the kind of climate that is created for the "yeast" of the children's relations to each other, the school is in a salient position to make an impact on the child's image of his human universe, to influence how safe and nondefensive he will feel with people, how benign and without hostility toward himself.

These complex goals make the educational process that actually transpires in the classroom a highly permeable one. For example, the subcultural mores of the family population influence the degree to which children can respond to friendly authority; the relative readiness of parents to relax from the concept of competitiveness as the primary motive for learning influences how reports and report cards are designed and used. More locally, within the school society itself, the quality of the teacher's relations to the children is part of an equation that contains the quality of the administrators' relations to the teachers; codes about noise, traditions for ordering multiple copies of a single text, accessibility of play space or running water, expectations of what assembly programs should be—all condition the degree to which a curriculum can be sensitive to and fulfill cognitive-emotional growth needs.

These factors do not exist as static context and should be weighed appropriately in any schemes for evaluation of the teacher. They interact with the teacher's influence which holds the highest potency and operates toward the achievement of goals in accordance with: (a) the kind of professional training he has had; (b) the array of interests, motives, and attitudes that he brings to his work as a teacher of children; and (c) the degree to which the integration of professional training and personal experience has prepared him for a psychodynamically conceived teaching role.

[2] See "Teacher Education in Mental Health: From the Viewpoint of an Educator," by Barbara Biber, Ph.D., and "From the Viewpoint of a Psychiatrist," by Viola W. Bernard, M.D., to be published as a section in a monograph by the American Orthopsychiatric Association.

Elaboration of how this preparation for teaching can be made relevant to the goals as postulated is not possible in this discussion.[3]

What Is "Good" in Teaching?

Instead, we turn to the question of how to assess, measure, evaluate teaching performance and to face the challenge of defining what is "good" as viewed in the perspective of the accepted goals. Unfortunately, meeting this challenge calls for a preliminary detour around three questions before any objective evaluation can proceed, whether as part of the practical life of promotion, self-criticism, staff development or part of a systematic search for reliably measurable criterion data.

1. *Can "good" be estimated in terms of effect on the pupils?* Granted that this is the ultimate measure of significance, it is nevertheless true that, when goals are conceived on a complex level, involving thought-processes as well as fact assimilation and subjective affect as well as adaptive behavior, it is not possible to short-cut intermediary factors and seek for direct measures of teacher-effectiveness in terms of effects on pupils.

For several reasons: the effects we are concerned with include subtle elements of personality for which measures are not readily available nor easily applicable, except by specially trained personnel; they require study of individuals and can only be superficially probed through group measures. Furthermore, it cannot be assumed that the teacher's influence in terms of these goals [4] will "take" over small units of time such as months or school terms nor that it will be free of the compounding influences of family, neighborhood and total school atmosphere. From this viewpoint, it may be possible, by controlling family, social-economic, and other differences, to study pupil growth as a measure of the cumulative effect of the educational experience, but not as an index of an individual teacher's effectiveness.[5]

Teachers, exposed to and accepting of these complex goals, have an important resolution to make concerning their own worth as teachers. They have to come to terms with the exhilaration they feel in instances of success and the defeat they experience where they seem to be failing. Some see themselves as the responsible, central figures in these outcomes; others feel themselves pushed, fatalistically, by forces outside themselves. Still others feel personal toward the successes, fatalistic toward the failures. Somewhere, inside the entanglement of observation and rationalization, the unmeasurable truth undoubtedly lies. Any increase in degree of awareness of complexity should be an asset to any teacher's wisdom and happiness.

2. *Is "good" a point, a range, a series of balances adapted to varying conditions?* Teaching is not a series of prescribed acts, to be measured good or bad in terms of frequency of occurrence or omission. It is a nexus of relationships taking shape from the way in which general goals and ideals are adapted to the particularity of given situations. In practical terms, this means that, for purposes of evaluation, each aspect of teaching needs to be considered in terms of a possible range within which "good"

[3] *Ibid.*

[4] The problems associated with pupil growth as a measure have been developed by William Rabinowitz and Robert M. W. Travers, "Problems of Defining and Assessing Teacher Effectiveness"—*Educational Theory*, Vol. III, No. 3, July 1953.

[5] Such a study is now in progress at the Bank Street College of Education under the title, "The Psychological Impact of School Experience."

teaching can fall. One of the scales we have used in a series of observations of teacher performance may serve best as illustration. This particular measurement was concerned with the extent to which the teacher imposes limits and restrictions on the children; five defined points are listed below:

a. Children's behavior is severely limited, beyond the reasonable requirements of their welfare

b. Requirements are abundant but aimed at practical ends, not "pure discipline"

c. Regulations are somewhat liberal and realistic; freedom is permitted in matters often subject to restriction

d. Children are expected to meet a few basic standards, but otherwise enjoy great freedom

e. Limits are both very scarce and very mild. The freely expressed impulses of the children are almost completely accepted.

In this scale, points "b," "c," and "d" represent variations of a function which could be considered within a range of "good" depending upon the situational context in which they exist. That this viewpoint does not reduce itself to an open-ended, relativistic position becomes clear in the definitions of points "a" and "e" which lie outside the range of acceptably good teaching.

3. *Can "good" be measured directly or inferred from a series of measurements?* In addition to the variation that is a function of adaptation to varying conditions, there is the variation within the person of the teacher himself, a factor whose importance is in proportion to the extent that teaching is psychologically conceived and recognized as subject to unconsciously derived fluctuations of preference, interest, tension or involvement. It is obviously futile, therefore, to measure any teacher against a fixed concept of a consistently ideal teacher. Most reliable evaluations can be expected where judgments of how "good" are made as an end-process of consideration based on varied measures of a single teacher and a judgment of the series of balances made by a teacher in adapting goals to situations.

These detour considerations yield a few guiding principles but, having rejected pupil growth as a suitable means of teacher evaluation, there still remain the key problems of *what* shall be taken as a measure of teaching and *how* this can be done. An alternate approach lies in devising means of characterizing a classroom in terms of its intrinsic learning processes and interpersonal relationships, thus arriving at an approximate measure of what is generated by the teacher as a climate of learning. Objective observations of teacher performance and children's response can serve as the core; additional methods, such as an interview, are necessary in order to place observed behavior in the context of attitudes and feelings upon which the meaning of behavior depends. To take this approach to evaluation opens one to the criticism of begging the ultimate question: how successfully is the teacher achieving goals with respect to the learning and growth of the children? In reply, one can only say that that large question had best be answered in several stages. In evaluating a teacher in terms of the learning climate he can create, one is assuming, with awareness, a relation between certain learning climates and ultimate effects on the children. This relation requires separate and independent testing-out to be undertaken under controlled research conditions.

In order to make concrete this approach to evaluation, it may be helpful to present a series of dimensions briefly defined, now in use in a study involving teacher performance as criterion data and based on the educational orientation implicit in

the goals presented at the beginning of this article. The data on which the characterization of each teacher is based include four observations, approximately one hour in length, made by four different observers using a pretested method of scaled observations and an individual interview, slightly more than an hour in length, covering a prepared content guide. The following dimensions are being used as a means of integrating observational and interview material.

- *Perception and acceptance of children.* Range of acceptance; kinds of child behavior accepted or rejected (liked, tolerated, disliked); degree of differentiation; bases of differentiation; extent to which acceptance is conditional, temporary.
- *Quality of relatedness to children.* Degree of empathic sensitivity to child-level of feeling, emotion, conflict; awareness of child level of thinking, capacity, impulse, etc.; spontaneous reaction to children; degree of ease; modes of communication; levels of connectedness.
- *Psychological understanding.* Context in which behavior is seen: developmental levels; cultural conditioning; social class influences, etc.; penetration beyond behavior level: motivational basis; multi-determined; etc.; degree of stereotypy in typing children or labeling behavior.
- *Approach to socialization.* Conception of socialization process; values and goals; nature of social process embedded in program organization; rules and regulations governing child interaction; provision for and response to individuality (spontaneity, originality, nonconformity).
- *Approach to learning.* Degree to which learning is motivated, interesting and child-participant; techniques for accomplishing above; degree to which learning consists of fact-assimilation, rote repetition, inquiry, independent thinking, individualized exploration.
- *Model of adult authority.* Nature and extent of restraints imposed; techniques of control; ways of dealing with infringements; forms of criticism; concept and quality of authority figure role.
- *Subjective relation to teaching.* Ideal teacher-figure; self-evaluation as a teacher; motivation for teaching; degree of ego-involvement; satisfaction-dissatisfaction with teaching experience; role as a teacher vis à vis parents.

The evaluation of teaching performance in terms of the dimensions briefly sketched above presents many challenging problems in research studies requiring quantification of measures, problems which must be faced once learning is conceived in terms of complex goals. The same scheme may have some direct value as a conceptual orientation to be used empirically in the continuous, everyday process of evaluation and self-evaluation of teaching.