

Stake R E. The countenance of educational evaluation. *Teach. Coll. Rec.* 68:523-40, 1967.

[Center for Instructional Research and Curriculum Evaluation, University of Illinois, IL]

What people perceive as the nature and quality of an educational program is determined in part by what an evaluator chooses to observe. In this paper I presented a matrix of eligible data and urged broader, more contextual description. [The SSC<sup>®</sup> indicates that this paper has been cited in over 130 publications.]

## Evaluation of Post-Sputnik Curriculum Reform

Robert E. Stake  
270 Education Building  
University of Illinois  
Champaign, IL 61820

January 12, 1990

I wrote my "countenance paper" because I was dismayed by the narrow selection of data being used for the formal evaluation of post-Sputnik curriculum reform efforts.

In 1965 the schools were getting a new array of curriculum packages. A new technical specialty, program evaluation, was charged to discover how good these instructional packages were. Fellow evaluation specialists had put forth a variety of methodological advice. Ben Bloom<sup>1,2</sup> spoke of evaluation as a special case of student testing. Jim Popham<sup>3</sup> advocated inquiries structured to behavioral objectives. For Don Campbell<sup>4</sup> the orientation was experimentation; for Lee Cronbach,<sup>5</sup> instructional development; for Dan Stufflebeam,<sup>6</sup> administrative decision making; and for Michael Scriven,<sup>7</sup> consumer service.

In 1964, as president of the American Educational Research Association (AERA), Cronbach named a committee to study the need for standards for the conduct of evaluation studies. Committee members were Nate Gage, Wells Hively, John Mayor, and me. We be-

came persuaded that it was a time not for standardization but for expansion, experimentation, and borrowing from other disciplines. We proposed that AERA sponsor a monograph series on curriculum evaluation to help explore design options. In Bloom's presidency it was begun.

The first volume of the series included Scriven's criteria for classifying evaluation situations. Early drafts of his monograph had been prepared for the Educational Consortium for Social Science, to which both Scriven and Cronbach were advisory committee members. In their meetings Scriven the philosopher enjoyed taking issue with Cronbach the psychologist, particularly disputing his priorities on formative evaluation, advocating instead that consumers be availed summative evaluation services, a scholarly endeavor that program evaluators could and should provide. In early drafts Scriven's opposition to Cronbach was sharply put. Cronbach appeared not aroused.

In 1965 Tom Hastings and I persuaded the two of them to come to Champaign-Urbana to clarify their differences. That evening Cronbach made a remark that puzzled me for years, yet presaged the direction my own work would take in the 1970s. He said something like, "What the evaluation field needs is a good social anthropologist." He was acknowledging the situational or cultural character of instructional programs and their resistance to sweeping generalization.

I didn't get the whole message, but I realized evaluators should provide contextual data. Few descriptive variables were absolutely essential; much was optional, the design depending on questions needing answers, which changed as time passed. And perhaps the whole was too easily shaped by the curiosities and talents of the evaluator. In keeping with the conclusions of the AERA committee, I was moved to explore emphases on local circumstances and program uniqueness. I began to write the countenance paper as the Champaign-Urbana discussions ended.

1. Bloom B S. *Toward a theory of testing which includes measurement-evaluation-assessment*. Los Angeles, CA: Center for the Study of Evaluation, 1968. 41 p. (Cited 190 times.)
2. -----, Toward a theory of testing which includes measurement-evaluation-assessment. (Wittrick M C & Wiley D E, eds.) *The evaluation of instruction: issues and problems*. New York: Holt, Rinehart & Winston, 1970. p. 25-69.
3. Popham W J. Objectives and instruction. (Stake R E, ed.) *Instructional objectives*. Chicago, IL: Rand McNally, 1969. p. 32-64. (Cited 10 times.)
4. Campbell D T. Reforms as experiments. *Amer. Psychol.* 24:409-29. 1969. (Cited 610 times.)
5. Cronbach L. Course improvement through evaluation. *Teach. Coll. Rec.* 64:672-83, 1963.
6. Stufflebeam D L. Evaluation as enlightenment for decision making. (Beatty W H, ed.) *Improving educational assessment and an inventory of measures of affective behavior*. Washington, DC: Association for Supervision and Curriculum Development, National Education Association, 1969. p. 41-73. (Cited 20 times.)
7. Scriven M. The methodology of evaluation. (Stake R E, ed.) *Perspectives of curriculum evaluation*. Chicago, IL: Rand McNally, 1967. p. 39-83. (Cited 95 times.)