

3

Responsive Evaluation

ROBERT STAKE

Center for Instructional Research and Curriculum Evaluation, University of Illinois at Champaign-Urbana, IL, USA

Responsive evaluation is an approach, a predisposition, to the evaluation of educational and other programs. Compared to most other approaches it draws attention to program activity, to program uniqueness, and to the social plurality of its people. This same predisposition toward merit and shortcoming can be built into or can be recognized in other approaches, such as a stakeholder evaluation or connoisseurship evaluation.

A responsive evaluation is a search and documentation of program quality. The essential feature of the approach is a responsiveness to key issues or problems, especially those recognized by people at the site. It is not particularly responsive to program theory or stated goals but more to stakeholder concerns. Its design usually develops slowly, with continuing adaptation of evaluation goal-setting and data-gathering in pace with the evaluators becoming well acquainted with the program and its contexts.

Issues are often taken as the “conceptual organizers” for the inquiry, rather than needs, objectives, hypotheses, or social and economic equations. Issues are organizational perplexities or complex social problems, regularly attendant to unexpected costs or side effects of program efforts. The term “issue” draws thinking toward the interactivity, particularity, and subjective valuing already felt by persons associated with the program. (Examples of issue questions: Are the eligibility criteria appropriate? Do these simulation exercises confuse the students about authoritative sources of information?) People involved in the program are concerned about one thing or another (or likely to become concerned). The evaluators inquire, negotiate, and select a few issues around which to organize the study.

The evaluators look for troubles and coping behavior. To become acquainted with a program’s issues, the evaluators usually observe its activities, interview those who have some role or stake in the program, and examine relevant documents. These are not necessarily the data-gathering methods for informing the interpretation of issues; but are needed for the initial planning and progressive focusing of the study. And even later, management of the study as a whole usually remains flexible – whether quantitative or qualitative data are gathered.

OBSERVATIONS AND JUDGMENTS

Directed toward discovery of merit and shortcoming in the program, responsive evaluation study recognizes multiple sources of valuing as well as multiple grounds. It is respectful of multiple, even sometimes contradictory, standards held by different individuals and groups.

Ultimately the evaluators describe the program activity, its issues, and make summary statements of program worth. But first they exchange descriptive data and interpretations with data givers, surrogate readers, and other evaluation specialists for recognizing misunderstanding and misrepresentation. In their reports they provide ample description of activities over time and personal viewing so that, with the reservations and best judgments of the evaluators, the report readers can make up their own minds about program quality.

There is a common misunderstanding that responsive evaluation requires naturalistic inquiry, case study, or qualitative methods. Not so. With the program staff, evaluation sponsors and others, the evaluators discuss alternative methods. Often the clients will want more emphasis on outcomes, and responsive evaluators press for more attention on the quality of processes. They negotiate. But knowing more about what different methods can accomplish, and what methods this evaluation “team” can do well, and being the ones to carry them out, the evaluators ultimately directly or indirectly decide what the methods will be. Preliminary emphasis often is on becoming acquainted with activity, especially for external evaluators, but also the history and social context of the program. The program philosophy may be phenomenological, participatory, instrumental, or in pursuit of accountability. Method depends partly on the situation. For it to be a good responsive evaluation the methods must fit the “here and now”, having potential for serving the evaluation needs of the various parties concerned.

Even so, it has been uncommon for a responsive evaluation study to emphasize the testing of students or other indicators of successful attainment of stated objectives. This is because such instrumentation has so often been found simplistic and inattentive to local circumstances. Available tests seldom provide comprehensive measures of the outcomes intended, even when stakeholders have grown used to using them. And even when possible, developing new tests and questionnaires right is very expensive. For good evaluation, test results have too often been disappointing – with educators, for example, probably justifiably believing that more was learned than showed up on the tests. With the responsive approach, tests often are used, but in a subordinate role. They are needed when it is clear that they actually can serve to inform about the quality of the program.

In most responsive evaluations, people are used more as sociological informants than as subjects. They are asked what they saw as well as what they felt. They are questioned not so much to see how they have changed but to indicate the changes they see.

SUBJECTIVITY AND PLURALISM

My first thoughts about how to evaluate programs were extensions of empirical social science and psychometrics, where depersonalization and objectivity were esteemed. As I have described elsewhere (Stake, 1998), in my efforts to evaluate curriculum reform efforts in the 1960s, I quickly found that neither those designs nor tests were getting data that answered enough of the important questions. Responsive evaluation was my response to “preordinate evaluation”, prior selection and final measurement of a few outcome criteria. Over the years I came to be comfortable with the idea that disciplining impressions and personal experience led to better understanding of merit and worth than using needs to identify improvement with strict controls on bias (Stake et al., 1997).

Case study, with the program as the case, became my preferred way of portraying the activity, the issues, and the personal relationships that reveal program quality. Not all who have a predilection for responsive evaluation use a case study format. Many evaluators do their work responsively without calling it that and some who do call their work responsive are not responsive to the same phenomena I am. There is no single meaning to the term.

Those who object to the responsive approach often do so on the ground that too much attention is given to subjective data, e.g., the testimony of participants or the judgments of students. For description of what is happening, the evaluators try (through triangulation and review panels) to show the credibility of observations and soundness of interpretations. Part of the program description, of course especially that about the worth of the program, is revealed in how people subjectively perceive what is going on. Placing value on the program is not seen as an act separate from experiencing it.

The researchers' own perceptions too are recognized as subjective, in choosing what to observe, in observing, and in reporting the observations. One tries in responsive evaluation to make those value commitments more recognizable. Issues, e.g., the importance of a professional development ethic, are not avoided because they are inextricably subjective. When reporting, care is taken to illuminate the subjectivity of data and interpretations.

Objection to a responsive approach is also expressed in the belief that a single authority, e.g., the program staff, the funding agency or the research community, should specify the key questions. Those questions often are worthy of study, but in program evaluation for public use, never exclusively. There is general expectation that if a program is evaluated, a wide array of important concerns will be considered. Embezzlement, racial discrimination, inconsistency in philosophy, and thwarting of creativity may be unmentioned in the contract and not found in the evaluators' expertise, but some sensitivity to all such shortcomings belong within the evaluation expectation, and the responsive evaluator at least tries not to be blind to them.

Further, it is recognized that evaluation studies are administratively prescribed, not only to gain understanding and inform decision-making but also to legitimize and protect administrative and program operations from criticism,

especially during the evaluation period. And still further, evaluation requirements are sometimes made more for the purpose of promulgating hoped-for standards than for seeing if they are being attained. Responsive evaluators expect to be working in political, competitive, and self-serving situations and the better ones expose the meanness they find.

By seeking out stakeholder issues, responsive evaluators try to see how political and commercial efforts extend control over education and social service. They are not automatically in favor of activist and legitimate reform efforts, but they tend to feature the issues they raise. Responsive evaluation was not conceived as an instrument of reform. Some activists find it democratic; others find it too conservative (Shadish, Cook, & Leviton, 1991). It has been used to serve the diverse people most affected personally and culturally by the program at hand – though it regularly produces some findings they do not like.

ORGANIZING AND REPORTING

The feedback from responsive evaluation studies is expected to be in format and language attractive and comprehensible to the various groups, responsive to their needs. Thus, even at the risk of catering, different reports or presentations may be prepared for different groups. Narrative portrayals, story telling, and verbatim testimony will be appropriate for some; data banks and regression analyses for others. Obviously the budget will not allow everything, so these different communications have to be considered early in the work.

Responsive evaluation is not participatory evaluation, but it is organized partly around stakeholder concerns and it is not uncommon for responsive evaluation feedback to occur early and throughout the evaluation period. Representatives of the prospective audience of readers should have directly or indirectly helped shape the list of issues to be pursued. Along the way, the evaluator may ask, “Is this pertinent?” and “And is this evidence of success?” and might, based on the answer, change priorities of inquiry.

Responsive evaluation has been useful during formative evaluation when the staff needs more formal ways of monitoring the program, when no one is sure what the next problems will be. It has been useful in summative evaluation when audiences want an understanding of a program’s activities, its strengths and shortcomings and when the evaluators feel that it is their responsibility to provide a vicarious experience. Such experience is seen as important if the readers of the report are to be able to determine the relevance of the findings to their own sense of program worth.

As analyzed by Ernest House (1980, p. 60) responsive evaluation will sometimes be found to be “intuitive” or indeed subjective, closer sometimes to literary criticism, Elliot Eisner’s connoisseurship, or Michael Scriven’s *modus operandi* evaluation than to the more traditional social science designs. When the public is seen as the client, responsive evaluation may be seen as “client centered”, as did Daniel Stufflebeam and Anthony Shinkfield (1985, p. 290). But usually it

differs from those approaches in the most essential feature, that of responding to the issues, language, contexts, and standards of an array of stakeholder groups.

When I proposed this “responsive evaluation” approach (at an evaluation conference at the Pedagogical Institute in Göteborg, Sweden, in 1974) I drew particularly upon the writings of Mike Atkin (1963); Lee Cronbach (1963); Jack Easley (1966); Stephen Kemmis (1976); Barry MacDonald (1976); and Malcolm Parlett and David Hamilton (1977). They spoke of the necessity of organizing the evaluation of programs around what was happening in classrooms, drawing more attention to what educators were doing and less attention to what students were doing. Later I reworked some of my ideas as I read Ernest House (1980); Egon Guba and Yvonna Lincoln (1985); Tom Schwandt (1989); and Linda Mabry (1998). Of course I was influenced by many who proposed other ways of evaluating programs.

It is difficult to tell from an evaluation report whether or not the study itself was “responsive.” A final report seldom reveals how issues were negotiated and how audiences were served. Examples of studies which were clearly intentionally responsive were those by Barry MacDonald (1982); Saville Kushner (1992); Anne McKee and Michael Watts (2000); Lou Smith and Paul Pohland (1974); and Robert Stake and Jack Easley (1979), indicated in the references below. My meta-evaluation, *Quieting Reform* (1986), also took the responsive approach.

REFERENCES

- Atkin, J.M. (1963). Some evaluation problems in a course content improvement project. *Journal of Research in Science Teaching*, *1*, 129–132.
- Cronbach, L.J. (1963). Course improvement through evaluation. *Teachers College Record*, *64*, 672–683.
- Easley, J.A., Jr. (1966). *Evaluation problems of the UICSM curriculum project*. Paper presented at the National Seminar for Research in Vocational Education. Champaign, IL: University of Illinois.
- Greene, J.C. (1997). Evaluation as advocacy. *Evaluation Practice*, *18*, 25–35.
- Guba, E., & Lincoln, Y. (1981). *Effective evaluation*. San Francisco: Jossey-Bass.
- Hamilton, D., Jenkins, D., King, C., MacDonald, B., & Parlett, M. (Eds.). (1977). *Beyond the numbers game*. London: Macmillan.
- House, E.R. (1980). *Evaluating with validity*. Beverly Hills, CA: Sage.
- Kushner, S. (1992). *A musical education: Innovation in the Conservatoire*. Victoria, Australia: Deakin University Press.
- Mabry, L. (1998). *Portofios plus: A critical guide to alternative assessment*. Newbury Park, CA: Corwin Press.
- MacDonald, B. (1976). Evaluation and the control of education. In D.A. Tawney (Ed.) *Curriculum evaluation today: Trends and implications*. London: Falmer.
- MacDonald, B., & Kushner, S. (Eds.). (1982). *Bread and dreams, CARE*. Norwich, England: University of East Anglia.
- McKee, A., & Watts, M. (2000). *Protecting Space? The Case of Practice and Professional Development Plans*. Norwich, England: Centre of Applied Research in Education, University of East Anglia.
- Parlett, M., & Hamilton, D. (1977). Evaluation as illumination: A new approach to the study of innovative programmes. In D. Hamilton, D. Jenkins, C. King, B. MacDonald, & M. Parlett (Eds.), *Beyond the numbers game*. London: Macmillan.
- Schwandt, T.A. (1989). Recapturing moral discourse in evaluation. *Educational Researcher*, *18*(8), 11–16.
- Shadish, W.R., Cook, T.D., & Leviton, L.C. (1991). *Foundations of program evaluation*. Newbury Park, CA: Sage.