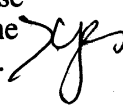


WHEN POLICY IS MERELY PROMOTION, BY WHAT ETHIC LIVES AN EVALUATOR?¹

Robert Stake, University of Illinois

This paper on the ethics of program evaluation reflects an increasing American departure from the traditional expectations of social science, with increasing postmodern doubt and cynicism--perhaps uncharacteristic of educational research in Israel.

Program evaluation is a technological subdivision of educational research. It includes objective and subjective inquiry disciplined search for program quality. Although utility, i.e., usefulness, is not fundamental to the concept of evaluation, contracts for program evaluation stipulate that the findings are to be useful for managing the program and representing it to various constituencies.

Some theories of evaluation, such as Dan Stufflebeam's and Michael Patton's, put utility as the central focus. Reports from their kind of evaluation devote *much less* attention to program quality *and more* to decision making and capacity building. Evaluators differ in their commitments as to whom they will be useful, with those following Lee Cronbach most attentive to program developers, those following Michael Scriven most attentive to the concerns of consumers, and those following Ernie House attending to the concerns of the least privileged. However, the *practice* by most of the evaluators I know is to attend to the interests of the party that pays for the evaluation. 

Ethics in program evaluation. The rationale for all approaches to program evaluation presumes an audience interested in information and interpretation. But the reality is that most formal, funded, evaluation studies are done so that an institution or agency or administrator will be recognized as prudent and accountable in calling for the study, or so that the findings will substantiate and support action anticipated, or already taken, such as requesting further funding—both examples of promotionalism. Program evaluation, for at least twenty years, has been an instrument of institutional promotion more than an effort at understanding its processes and quality.

The theory and practice of evaluation are of little value unless we can count on vigorous ethical behavior by evaluators. My plan today is to describe two evaluation studies from CIRCE, our evaluation center at the University of Illinois, studies in which we have faced ethical problems--particularly noting the atmosphere of promotion and advocacy.

Example 1. Chicago Teachers Academy. My CIRCE colleagues and I finally submitted the long-delayed report evaluating the Chicago Teachers Academy, a free-standing professional development institution devoted to improved teaching of children in the urban schools. In this and previous years, our reports told of the Academy's sophisticated theories of pedagogy and curriculum on which are based their sequenced workshops for teachers from dozens of participating schools. With data from field studies, we described improvements in teaching in some schools. We told of the Academy School Improvement Unit's fine efforts at helping schools define and work-at their own organizational problems. Also as to quality, we found this Unit so strongly committed to helping the schools solve their problems themselves that they gave, we believed, too little help in examining problem-solving efforts at other schools and too little acquaintance with the "institutional change" literature. Our *draft report* went into detail on activities at the participating schools.

¹ Paper presented at the University of Tel Aviv, January 29, 1998.

Ever since its beginning almost ten years ago, the Academy has had uncertain financial support. Proud of its independence, self-centered; it consequently gained little cross-institutional support. We saw the Academy as a greatly-needed school-assistance institution in its city, doing a specialized job no others were doing and without a potential successor in sight. We became acquainted with its mistakes, inefficiencies, and political indulgences. Admirers pointed to its virtues, particularly its commitment to experiential education and on-site, personal familiarity with individual teachers in affiliated schools. In about a dozen case studies, we looked at those virtues in detail. We found the Academy, clearly, a positive force for the continuing professional development of a large number of teachers.

Ideology. As evaluators of the Chicago Academy, and other evaluands, we at CIRCE work to find and understand quality, quality of the institution, quality of individual performance. We offer little assistance in remediation. We do not think of ourselves as collaborators in redevelopment. Directly and indirectly, we promise to help the Academy staff understand the quality of its operations. We provide details of merit and shortcoming. We interpret and discuss the issues.

We recognize that we evaluators have ideological positions on many of the issues. For example, at CIRCE we share a widespread view that in-service training strategies that depend on teachers and schools volunteering to participate poorly engage the teachers most needing assistance. And we, more or less alone, believe that standards for schools set by central authorities, however popular and legitimate, divert teachers from their own efforts at improving their work.

Many of our persuasions have not been substantiated by research or extensive professional experience. Some educators and evaluators have opposing views. By allowing these biases to color our accounts, we express a parochial ideology of quality which, of course, shapes our assertions of program merit. Such biases in professional judgment have long troubled some of our colleagues, particularly Michael Scriven, but we have resisted Michael's summons (1994) to make the processes rational, choosing to stick with an interpretive approach to the resolution of value questions. Unfortunately, some of our presumptions as to quality are deeply assumed and seldom sufficiently declared. Still, whether called bias or the exercise of disciplined value commitment, ideological interpretation in program evaluation is not a violation of ethics.

Expectations of nondisclosure and advocacy. Our draft of last year's evaluation of the Chicago Academy remained unfinished more than six months. One standoff was our claim that the Academy's newly developed training in matters of "assessment" was short-sighted, that the workshop developer was not using good judgment in his choice of activities and resources. The Director suggested we remove that section on the grounds that "their enemies" would use it against them. We failed to see that they had enemies, but we respected their political sensitivity. With no permanent funding, the Academy had to seek partial support by proposing a move into whatever training might be a current government priority. Clearly, for every new funding the Academy sought, older agencies saw encroachment on their territories. ✓

When resources drop below a certain level, the Academy *will* close. We worried about hastening its demise—but we were not persuaded to drop our criticism of assessment training. We felt that we had already prepared a report somewhat more supportive than fully warranted, and we did not remove the challenged section. Oh, yes, we had faced the question before, and sometimes *had* compromised, sometimes on

the Director's objection, but more often because, by ourselves, we decided that a full statement of the negative findings could be unduly hurtful.

Thus, we did not fully disclose findings we thought unnecessarily hurtful, and we refused to withhold some findings our client felt clearly hurtful. Forwarded to the client, the report remained in draft form months after its delivery date.

It is a generic dilemma of disclosure, to some degree accompanying every evaluation contract. When we contract for evaluation with program officers, they show distressingly little curiosity about the workings of their programs. Most expect us to become advocates for the good we find and discrete about the shortcomings. All the world fixates upon the bad, even in a sea of good. Clients seek hope in helping their programs survive. Yet sooner or later, those officials point to the evaluation report as evidence of their integrity or productivity or worthiness for funding. So, we have an obligation, partly to audiences we will never know, to represent the quality of the program as accurately as we can.

Codes of ethics. Do codes of ethics help? Even when accompanied with situational detail, they give us *broad* categories and *general* principles. Their examples are chosen for their generalizability. The most widely known code for our work is the Joint Committee's *Program Evaluation Standards*. In the box below, I have identified the four categories of its standards: Utility, Feasibility, Propriety, and Accuracy.

Joint Committee on Standards *Program Evaluation Standards*

Utility. The utility standards are intended to ensure that an evaluation will serve the information needs of intended users.

Feasibility. The feasibility standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.

Propriety. The propriety standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.

Accuracy. The accuracy standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.

The ethical aspects of our work fall largely in the Propriety category, detailed below.

Joint Standards of Propriety for Evaluators

P1. **Service organization.** Evaluations should be designed to assist organizations to address and effectively serve the needs of the full range of targeted participants.

P2. **Formal agreements.** Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing so that these parties are obligated to adhere to all conditions of the agreement or formally to renegotiate it.

P3. **Rights of human subjects.** Evaluations should be designed and conducted to respect and protect the rights and welfare of human subjects.

P4. **Human interactions.** Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation so that participants are not threatened or harmed.

P5. **Complete and fair assessment.** The evaluation should be complete and fair in its examination and recording of strengths and weaknesses of the program being evaluated so that strengths can be built upon and problem areas addressed.

P6. **Disclosure of findings.** The formal parties to an evaluation should ensure that the full set of evaluation findings along with pertinent limitations are made accessible to the persons affected by the evaluation and any others with expressed legal rights to receive the results.

P7. **Conflict of interest.** Conflict of interest should be dealt with openly and honestly, so that it does not compromise the evaluation processes and results.

