

# Evaluation: Purposes and Levels

## Evaluation: Objectifs et niveaux

GEORGE L. GEIS\*

Although they work in different worlds, both marriage counsellors and statistical consultants are likely to complain to their clients, "Why didn't you come to me earlier?" Consultants in educational evaluation share the problem: the client typically comes for advice and assistance long after an evaluation plan has been developed and implemented. It is as if the client arrived with a shopping bag of macaroni, turnips, cornflakes, and a soup chicken and demanded that the magical recipes of the evaluation expert produce a gourmet meal.

With the rapidly increasing popularity of the evaluation of courses, programs, and teachers, evaluators are under increasing pressure to communicate to the users a clearer picture of how evaluations ought to be planned and carried out. This paper deals with two concepts of educational evaluation and indicates how they illustrate the need for early planning.

### Purposes

Evaluations are purposeful (except for some esoteric ones that are carried out by autistic academics for their own sake) and occur when decisions are to be made. The evaluation and the data on which it is based are seen as serving an important input into the decision-making process. This is not to say that decisions are not made with explicit formal evaluation. Nor to suggest that decision-makers always adequately take into account the results of formal evaluations. However when an evaluation is undertaken, usually it is related to specific decisions that are to be made further along in the time line.

The purposes however, are rarely explicated at the beginning of the evaluation planning. Indeed, clear explication is often actively avoided; perhaps it is seen as too time consuming. Or in some cases such an explication might cause perturbations in the system: ambiguity at least allows one to get on with the job.

I would suggest that under these circumstances the job is probably not worth doing. Early explication is an important activity; one of *its* purposes is to alert all of those who will be concerned with the evaluation to the kinds of decisions that are to be made. Instead of disguising the connections between evaluation and decision options, the profes-

\*Director of the Centre for Learning and Development, and Associate Professor in Education, McGill University.

sional evaluator should take an active role in explicating the relationships. Let those who will object or argue do so early and let them direct their energies at the decision-making activity itself. Otherwise, there is a high probability that later on they will work to suppress, attack, and possibly pervert the evaluation effort, results, or role in decision-making.

Furthermore, linking the evaluation plan to decision options provides the information that is needed in order to select methods of evaluation and appropriate data treatment, and to estimate the resources that will be involved in the evaluation effort.

The evaluator himself may be somewhat naive about the relationship between the evaluation effort and decisions; for instance, decisions often have to do with the allocation or re-allocation of resources, can affect role and status, or may jeopardize jobs. Consequently, like it or not, evaluation is going to be in the context of critical, meaningful, political, and personal decisions. To view evaluation as an isolated, impersonal, and "objective" process is a serious mistake. Such an attitude may well be one of the important reasons why evaluation efforts rarely have any impact on decision-making. Decisions are made by individuals with needs and values and if those needs and values are not expressed early in the evaluation planning, they will be expressed in the decision-making process. An educational evaluation system under development at McGill University not only encourages the client to explicate the decisions at which the evaluation is aimed but also suggests simulating decision-making in the face of dramatically different data. For example: "What actions would you take if the student responses on this questionnaire item were preponderantly negative?"

*Kinds of Decisions.* Kinds of decisions and, consequently, purposes of evaluation, may be broadly classified as *summative* and *formative*. Summative evaluations eventuate in binary decisions. The decision to promote or not to promote a professor, to offer tenure or not, to pass or fail a student, to continue or drop a program, to buy or not to buy a particular textbook, or to implement a certain innovative teaching method would be summative decisions. Commonly these kinds of decisions are selective; they presume a sizeable field of items or people from which to choose; and they are often heavily influenced by the economies of the situation.

Formative evaluation and the consequent formative decisions, on the other hand, usually accept the item or person undergoing evaluation as a continuing given in the system. Formative activities look to improve the existing system through development and remediation rather than through rejection and selection procedures.

All of the people who will be involved in the proposed evaluation ought to be at least aware of, and hopefully supportive of, the kind of evaluation. Far more serious than the delay in starting the evaluations, because of the exercise in explicating purposes and kind, is the confrontation that can occur post-evaluation when the purpose becomes clear and threatening. Thus an evaluation that is passed off as one that is designed to provide useful information to the teacher who wishes to improve his course (e.g., data gathered via a student questionnaire) may in fact produce information that is used by administrators in decisions about promotion and tenure. Had the evaluation been seen as summative early in planning there would have been the opportunity for venting the confrontation that is likely to occur later in the absence of such explication; the evaluators might have developed a better, fairer, more comprehensive plan for the evaluation; and, perhaps, discussion would have opened up on general issues of information-gathering for administrative pur-

### 3 Evaluation: Purposes and Levels

poses. These issues otherwise will lay latently underneath the surface until the decision is about to be made or has been made. Conversely, if the purpose is clearly seen as formative, and appropriate safeguards against administrative use are set up, then one might expect more cooperation from those being evaluated; again, more appropriate instruments of evaluation can be employed; and the issues of resources to be employed after formative data are gathered can be raised early in the game.

*Effectiveness Versus Value.* A second aspect of purpose also requires early explication. Frequently there is a confusion between evaluation in order to determine the effectiveness of the system in meeting certain goals and the evaluation of the *goals* themselves.

A typical scenario runs like this: Complaints about a course or program are articulated in terms of effectiveness. An evaluation plan is mounted to determine how effective the system really is. Data are gathered which show, let us say, that the particular program under study has the impact that is claimed for it (e.g., students acquire and exhibit the outcome goals of the course). At that point the critic comments that, regardless of the efficiency of the system, there is really no point in teaching those things anyway — essentially an *ex post facto* statement of purpose of the evaluation.

Goals should (and, incidentally, can) be subjected to empirical evaluation. But the evaluation of goals ought to *precede* the evaluation of the effectiveness in meeting those goals. Explicating goals before evaluating the attempts to meet those goals will be discussed in more detail shortly. The point being made here is that evaluations which are aimed at determining the effectiveness of an implementation system are unlikely to cast light on the value of the goals to which the system is directed.

#### Levels

Educational systems can be looked at macroscopically or microscopically. The evaluation planners ought to be aware of the particular level the evaluation is aimed and what constraints and opportunities are involved in using data derived from a particular level.

At the broadest level, one would be dealing with the university system (and usually with the particular university). Evaluation may be carried out of a summative or formative sort and the evaluation may be aimed at the explicating and valuing the goals of the university as well as determining the effectiveness of the university meeting these goals. Moving to a more microscopic level, any of these evaluation purposes might be directed at schools, departments, or programs. A still finer magnification would focus attention on the course. And finally particular elements of a course could be examined. Thus, to take an example through all the steps of magnification, one might start by asking whether the university values professional preparation and training as one of its goals. Then one could examine the various programs offered, explicating the goals of each program and determining the effectiveness of each program or school in meeting their own goals. The particular courses which make up a program can be examined as to value and effectiveness. And finally, the elements within courses can be so examined. It is at this most microscopic level that one could ask questions about particular textbooks, teachers, methods of grading and the like.

*Evaluating Teachers or Teaching.* This logical analysis of how one would evaluate education is contaminated in a sense by our almost obsessive focus not upon levels, goals, or

effectiveness, but upon professors. Usually educational evaluation quickly is distilled down to the evaluation of teachers. Purposes and levels become confused and resultant evaluations are often unfair (and are often ignored).

The confusion in part stems from the lack of fit between the model described so far in this paper and the actual roles that professors play on a typical campus. The professor often sets the goals for his or her own course, partakes in the setting of goals for the university and the development of program goals. He or she also is the administrator of instruction and often the major instrument of instruction. And further he or she has roles other than, and sometimes in conflict with, his or her role as a teacher and, rightly insists upon consideration of them in any overall evaluation of him/her as professor.

The evaluator must be clear about his or her own and the client's understanding with regard to this issue. If the evaluator speaks of the evaluation of education having in mind the evaluation of programs and courses but the listener interprets "education" to be almost isomorphic with "professor" there is likely to be confusion, misunderstanding and even injustice. From the outset the distinction between the evaluation of teaching and of the teacher should be made clear. Furthermore, if the teacher is to be evaluated as any other component in the instructional system might be (e.g., a textbook, film series), then the *particular* role and purposes must be isolated and explicated in detail if the evaluation data are to be accepted and made use of. On the other hand if the teacher is to be evaluated in regard to many roles (e.g., content expert, organizer, classroom instructor) then care must be taken to gather adequate information on each of these facets. And, of course, the evaluation of "teacher" as professor will bring in additional aspects of an academic's life on campus.

This paper is not intended as a full discussion of the issues raised but merely to bring them to the forefront, to emphasize the critical importance of early explication of fundamental issues in any evaluation. Such explication is critically important if the evaluation is to be carried out fairly and without bias and if it is to have any impact upon the decision-making which is designed to aid.

Evaluators cannot help but be stimulated and excited by the rush on most campuses to join the evaluation bandwagon. Paradoxically, it is more incumbent upon them than ever before to move slowly and carefully and to raise the awareness of their clients before acceding to this long sought after, and well deserved, recognition of their talent and importance.