

Evaluation Research and National Social Policy: An Academic Practitioner's Perspective

RONALD G. EHRENBERG

Society has limited resources and many competing uses for them. I therefore take it as being an almost obvious proposition that at any point in time policy makers should strive to maximize the social benefits produced by the available funds they have to spend. This proposition implies that evaluation research should be undertaken either by or for government agencies. Policy makers need to know what benefits are being produced by each social program and the resource costs involved. They need to know which aspects of which programs are working and which programs need to be replaced.

Any evaluation study of a social program can have four possible outcomes, depending on the study's results. First, the study may be challenged by those people who are unhappy with, or feel threatened by, the study's conclusions. Because virtually any study has some methodological deficiencies, and because it is easier to find fault than it is to design and conduct an evaluation, this occurs frequently. Second, if the evaluation is negative, the program may be killed and the funds used elsewhere. Third, the program may be expanded or contracted in scale. Fourth, the program may be modified or redesigned. The option chosen depends on policy makers' values and prior benefits; all four outcomes are consistent, for example, with a negative conclusion about the program's impact. Although evaluation research may itself be value-free, the uses to which the research is put are clearly not.

If the need for evaluation research is so clear, one might question why so few evaluations are done in Washington and why evaluation research is given low priority. One reason is that the rewards accruing

to individuals who undertake evaluations are quite minimal. For example, at the highest level of the bureaucracy (e.g., the Assistant Secretary level and up), individuals are rewarded for their new initiatives, not for evaluating existing programs. The length of tenure of a typical high-level administrator is less than two years. To make a reputation in such a short time, the executive must do dramatic things that call attention to him or her, not simply monitor and evaluate existing programs. To increase the amount of evaluation research being conducted in Washington clearly requires bureaucrats to have longer tenures in their jobs and to be rewarded for their substantive long-term efforts—not for their flashy short-run initiatives.

The second reason evaluation research does not proliferate in Washington is that many individuals with long-term connections to particular programs feel threatened by the prospect of having their programs evaluated. An evaluation conceivably might adversely affect the level of funding of a program or suggest that the program be restructured. Individuals closely connected with a program often conclude that an evaluation can only do them harm and seek to avoid, like the plague, being subjected to evaluations.

Indeed, this suggests that evaluation studies should be undertaken by *outsiders* who have no vested interests in either continuing or killing a program. It also suggests the wisdom of conducting evaluation studies under the sponsorship of agencies other than the operating agency that is running a program.

There are at least three constraints on the evaluation research process that severely restrict the quality and objectivity of evaluations done in Washington. First, the agencies funding the evaluations, even if somewhat removed from the operation of programs, often hope that the evaluations will support their positions on policy issues. For example, during the Carter administration I served as a consultant on a project evaluating the displacement effects of the public employment program component of the Comprehensive Employment and Training Act (CETA) program. The project focused on estimating the extent to which state and local governments used CETA funds to hire employees whom they would have hired even in the absence of the funds.

This evaluation was undertaken by the Urban Institute, a reputable private nonprofit research organization in Washington. This project was not sponsored by the Employment and Training Administration, but rather by the Office of the Assistant Secretary for Policy, Evaluation and Research (ASPER) of the U.S. Department of Labor. Under the Ford administration this office tended to summarize the available

evidence on the merits and demerits of proposed changes in legislation or social programs, without taking strong advocacy positions. However, during the Carter administration, ASPER was transformed into a policy advocate and became the prime supporter for the inclusion of a large-scale public sector employment program in the welfare reform legislation then being considered.

Because of this change in the nature of the office, high-level administrators in ASPER had a vested interest in seeing that the results eventually obtained by the Urban Institute evaluation did not cast public employment programs in a negative light. Knowledge of this fact placed the ASPER staffers who were monitoring the evaluation in a difficult position, as their desire to maintain scientific objectivity potentially was tempered by their concern not to undercut their "bosses' baby." In the main, in this case they were able to maintain their objectivity, although after at least one meeting, those of us involved in the evaluation came away convinced they had been over-critical of our methodological approach *because* we found significant displacement effects.

Independent of the merits of the public employment program, agencies sponsoring evaluations should not play such an asymmetric role. The research methodology should be critically analyzed by the sponsoring agency, regardless of whether the results of the study support the agency's prior position. More importantly, this example also suggests that evaluation components of the cabinet departments should be divorced from the policy advocate components. In particular, there should be an office in the Department of Labor that is free to conduct long-run program evaluations without having to get involved in short-run policy advocacy.

The second constraint on the evaluation process is time itself. Evaluations typically have short horizons (one year or less in length) because policy makers want results in time to utilize them during their tenure in office. (Recall that high-level policy makers tend to turn over rapidly.) This time constraint prevents many studies from being given careful consideration during the design stage, which is necessary if a methodologically sound study is to be conducted. Another time constraint relates to the manner in which proposals for evaluations are solicited. The typical request for proposals (RFP) is issued with a three-week deadline, preventing academics from bidding on the contract (does this sound like sour grapes?) and precluding most consulting companies from seriously considering what the optimal method of analyzing the problem would be, thereby further limiting the likely usefulness of the studies.

The third constraint on the evaluation process is that most evaluators are in the evaluation business on a long-run basis and consequently are interested in obtaining future evaluation contracts. They are reluctant to submit analyses whose results are in direct conflict with what they know policy makers want. It is hard for investigators to maintain complete objectivity when they know that reporting certain results will increase (or decrease) the probability that they will obtain future contracts from the same funding agency. That the potential for such a conflict of interests exists was obvious to me in the course of my participation in the Urban Institute's CETA evaluation. This problem again suggests that it would be wise to separate the evaluation office in a cabinet agency from the office that is in charge of advocating policy reform. Evaluation should be done in a neutral environment as much as possible, and the evaluator should not be forced to respond to the wishes of a policy advocate.

Independent of how we structure the process by which evaluation research is conducted in government, one might ask what the role of evaluation research is in public policy decisions. Evaluation research seeks to provide information about the benefits and costs of alternative programs and legislation. It is naive, however, to think that such information will dictate a policy decision in any particular situation or typically even play a major role in the decision for at least two reasons.

On the one hand, evaluations tend to be narrower in scope than what the sponsors of the studies (usually government agencies) really want. What is desired is to know whether a particular program is good or bad, but what is offered typically is an estimate of one of the program's many effects. For example, literally hundreds of econometric studies have focused on estimating the adverse effects of minimum wage legislation on employment. However, only a handful have looked at questions such as, *Is the legislation complied with?* or *Do minimum wage laws reduce the incidence of poverty?*—questions that surely must be addressed in any complete evaluation.

On the other hand, policy makers' decisions are based on a complex set of factors that are not restricted to the way the program is operating, but often hinge on political considerations. Because of this, all one can hope is that evaluation research will provide some information on the workings of the programs and that this information will allow such decision makers, who are at the margin, to make more informed decisions. Anyone who conducts evaluation research with the anticipation that his or her evaluation study will, *and should*, automatically be transferred into policy recommendations will perpetually be frustrated and should probably find an alternative occupation.

Some distinguished academics and policy makers, such as Harvard's John Dunlop, a former Secretary of Labor, have argued that the benefit/cost ratio for evaluation research is zero. The argument proceeds along the following lines: Evaluation research uses resources, but it rarely influences decisions, so its benefits must be close to zero. As such, the benefit/cost ratio for conducting evaluations is close to zero, and it is a waste of resources to undertake the studies. This evaluation of evaluation research suggests that society needs less of it rather than more.

I draw the opposite conclusion. Rather than eliminating such studies, one can increase their benefit/cost ratio by increasing their benefits via utilizing them more in policy decision. It is incumbent upon academic economists who accept policy positions in the federal government, as Dunlop did, to increase the extent of rationality in the policy decision process by making more use of these studies.

COMPREHENSION QUESTIONS

1. According to the article, evaluation research should be undertaken to:
 - a. enable program supervisors to hire the right people to implement the program
 - b. enable program supervisors to be changed without affecting the program itself
 - c. enable policy makers to conduct cost/benefit analysis
 - d. facilitate proper record keeping
2. Evaluation research is:
 - a. value-free, as are the uses to which it is put
 - b. not value-free, but the uses to which it is put *are* value-free
 - c. value-free, but the uses to which it is put are *not* value-free
 - d. not value-free, nor are the uses to which it is put
3. The length of tenure of a typical high-level administrator is less than:
 - a. nine months
 - b. one year
 - c. eighteen months
 - d. two years
4. Evaluation studies should be conducted by:
 - a. the director of the program under evaluation
 - b. the person who hired the program director
 - c. an outside agency
 - d. none of the above
5. The author discusses which set of problems confronting those conducting evaluation research:
 - a. time, bias
 - b. bias, cost
 - c. cost, time
 - d. time, cost, and bias

SOCIAL SCIENCES

6. Sponsors of studies often find evaluations to be:
 - a. too broad in scope
 - b. too narrow in scope
 - c. too expensive
 - d. too time-consuming
7. The author believes that the benefit/cost ratio for evaluation research:
 - a. is higher than one would rightfully expect
 - b. is zero
 - c. should be increased through greater utilization
 - d. should be decreased by eliminating waste and duplication
8. Concerning his own bias, the author:
 - a. made no attempt to hide it
 - b. unsuccessfully tried to hide it
 - c. successfully hid it
 - d. eliminated it altogether
9. The author's structural organization of his subject is best characterized as:
 - a. chronological and spatial
 - b. descriptive enumeration
 - c. comparison and contrast
 - d. illustration and quotation
10. The main idea of the article is:
 - a. Benefits of evaluation research studies should be maximized through greater utilization when planning national social policy.
 - b. The federal government spends too much time and money on evaluation research studies, which, all too often, are unnecessary.
 - c. Evaluation studies should be conducted by academicians who have no vested interest in either maintaining or eliminating the program under evaluation.
 - d. Although national social policy has changed greatly under the last three presidents, it has not been improved.

ANALYSIS QUESTIONS

1. Ehrenberg refers to himself as an academic practitioner. How would his topic be approached by a cabinet official? a clerk in a social security office? a government economist?
2. Ehrenberg maintains that evaluation must be done properly, and without bias, to be both fair and useful. Apply this view to faculty grading of students. To student evaluation of faculty. To evaluation by one's peers on a jury.
3. How can the results of evaluation studies be translated into public policy?
4. If you headed a delegation to overhaul the social security system, how would you proceed?