49585

REPORTS OF THE THREE WORKING PANEL CHAIRPERSONS   NOVEMBER 19, 1976

I. WORKING PANEL III: IMPROVING THE
    UTILIZATION OF EVALUATION FINDINGS

                    CHAIRMAN: BLAIR G. EWING, Acting Deputy Director,
                                National Institute of Law Enforcement
                                and Criminal Justice
                                Law Enforcement Assistance Administration

MS. CHELIMSKY:

        Can we convene now for the final reports of the panels? We
have had, as you know, three groups meditating and reflecting on
various aspects of our evaluation problems: Working Panels I, II and
III. I had naturally thought we would start with Panel I and go
through to Panel III, but given the kinds of issues which were, it
seems, actually examined by the panels, it now seems more logical
to reverse the order and start with Panel III. In this way, we can
examine what the various panels have had to say, first, about users
and conditions for use, second, about evaluation criteria and
their substance, and finally, about evaluator/agency working rela-
tionships. Do you want to start then, Blair?

MR. EWING:

        We are the panel on improving the utilization of evaluation
findings. In my introduction to panel discussions yesterday after-
noon, I said that it seemed to me that there were multiple uses of
evaluation findings, ranging from program development to resource
allocation, to killing programs, to covering various parts of adminis-
trators' anatomies, to planning, to the development of further
research and further evaluation, to budget justification, et cetera.
There were also a very large number of audiences within agencies for
evaluation findings and these audiences could range from the program
managers themselves (whose programs are being evaluated) to the
planners, researchers and evaluators in the agency, to the top

                                333

management of that agency. I also talked briefly about some of the conditions that I saw as being essential for the use of evaluation results by agencies and then our panel began an enormously lively and spirited discussion which reflected quite a sum of experience and many discordant viewpoints.

We did not have time to address adequately all the topics on our agenda, perhaps because of some lack of consensus among us (although there was, in fact, some agreement), or perhaps because we were so preoccupied with those we did discuss in depth. We focused on three major aspects of evaluation use and usability:

(1) The user or the audience for evaluation findings;

(2) The kinds of information needed by that user, that audience; and

(3) The conditions which stimulate or impede the use of evaluation findings by agencies.

First, Evaluation Users. We began by examining the question of who uses evaluation findings, and decided that although there are many potential users within a given agency, the primary audience would depend on who needed the evaluation, and on the evaluation's character and scope. Given that users are pluralistic (decision-makers sit at different levels) and that there are many possible conflicts among the information needs of different users, the panel agreed generally that evaluation must at least begin by addressing the needs of the person who asked for the evaluation. The character and scope of the evaluation are also important in determining the audience for the findings, in the sense that an evaluation of a small program's effi-ciency might have as its major users the program's manager and the agency budget officer; whereas the users of a comprehensive, large-scale evaluation of the effectiveness of an important agency program would be the agency's policy-makers, and then—in widening circles—the research community, OMB, GAO, the Congress, the press, the public.

334

It was pointed out also, and largely agreed, that the audience
for evaluation findings which concerned programs in the field, could
not be limited to people at the Federal level since those findings
needed to be implemented by state and local government people and
by the institutional practitioners (e.g., teachers, policemen,
nurses, etc.) whose work had been evaluated and whose efforts and
good will would be needed to improve the program. Panel members
felt that the Federal role in this area was to build knowledge and
that efforts are presently lacking to improve the local ability to
rank priorities or compare rationally among local programs as to
effectiveness and cost. There was some consensus that--in the words
of one participant--"When the Federal Government sponsors an evalua-
tion, that evaluation gets designed on the basis of assumptions
made by the Federal agency about what is of interest to locals. There
is little or no participation by locals in the evaluation design.
When the results come in, the Federal agency itself has difficulty in
understanding what they may mean (either to the Federal Government
or to state and local governments) and it has no strategy for com-
municating what they might mean to the local practitioners who are
intimately concerned." Finally, the point was made that, where
federal initiatives at the local level are concerned, there does not
seem to be much point in doing evaluations of "demonstration" pro-
grams unless there is some commitment on the part of local people
to institutionalize. In effect, if locals don't intend to continue
a project, their need for evaluation findings would appear to be some-
what diminished. As one panel member put it, "The Federal Government
has little leverage to ensure improvement at the local level, no
matter how good the evaluation."

Possible conflicts among the needs of evaluation users was dis-
cussed at length. We recapitulated some of the Agency Perspectives
Panel discussion by examining the public interest versus the agency

interest, Congressional and OMB oversight needs versus agency needs; we contrasted the Federal policy-maker with the local practitioner or implementer, and the Executive Branch generally, with the Legislative. Professor Martinson told us that "the fundamental function of evaluation, like other forms of social science, is to enlighten the public as to whether or not the agencies to which the public pays taxes is using that money properly." He felt that if that interfered with what he called "purely symbolic activity snugly ensconced in an agency," well, then so much the better. Most of the rest of the panel, however, felt that our panel was dealing with agency use of evaluation findings and that the users we should consider, therefore, had to be primarily the agency managers who had asked for the evaluation and/or needed the information it could furnish. One of our panel members (who represented a Federal agency) made the point that Executive Branch policy-makers cannot change important agency policy without Congressional assent; yet often, an effort to change agency policy because of feasibility, or cost/effectiveness considerations, runs up against Congressional attention to special, powerful constituent groups. Therefore, it is wise as well, to build in, early on, both Congressional knowledge and use of agency evaluation.

Second: User Information Needs. It seems a natural assumption that Federal agencies would be more likely to use evaluation findings which produce information needed by agency managers. From there, it seems only a small step to ask the decision-maker who called for the evaluation what he expects from it, what it is he needs to know. Our panel members did indeed agree that the question of whether or not evaluation is used by agencies does depend in large measure on whether the right questions have been asked. The problem is that it is often very difficult to find out what these "right questions" are, especially in evaluations of complex programs.

To begin with, all questions are not answerable, so the first problem is to find the three (or so) questions which can feasibly be addressed by the evaluation and which are important to the decision-maker. But, as one panel member pointed out, many decision-makers do not themselves always know the "right" questions to ask, and here the panel felt it might be a useful learning experience for policy people to be involved in evaluation planning. "The real issue" said an evaluator member of our panel "is training managerial people to understand the limits of evaluation," how it can be used, what can be asked of it.

Here we had a split in our panel. Some people felt that the way to find out the right questions was through direct interaction between evaluators and agency managers, that the latter don't need to understand the limits of evaluation. They pointed out that perhaps decision-makers do not need to ask questions better because there is too much lack of consensus in social program areas. "What decision-makers are really interested in," said one panel member (a decision-maker himself), "is in keeping the system operating and stable, in not letting the temperature go too high or too low. He doesn't want to transgress boundaries, he wants to know when it's too hot and when it's too cold, and whether the thermostat moves quicker in a heating or a cooling system. He wants to know whether there is money waste, and he wants to know whether there is any visible achievement, or any visible failure to achieve. Those are the 'right questions' for him."

Other panel members pointed out that the "right questions" depended upon the type of evaluation envisaged, that many agencies use evaluation almost exclusively as a management tool and that questions of program achievement and effectiveness could rarely be

337

addressed by such evaluations. So that, to promote use and avoid disappointment, it becomes very important that decision-makers understand what questions can be asked of a particular evaluation and how this information obtained can then be used. Some members suggested that participation in evaluation planning might be a useful exercise for allowing agency managers to familiarize themselves with the possibilities and limitations of various evaluation strategies.

My own feeling is that an important problem in establishing what questions to ask is that it is very rare (at least in my experience) for managers to call for evaluation in order to improve planning and decision-making. The questions they ask, and what they want to know, is a function of why they asked for the evaluation in the first place. Usually they ask for evaluation:

- when they are stuck with a program they mistrust and want to cover themselves;
- when the program is in an enemy's province (evaluation is here used as an assassination instrument);
- when they don't understand a program and want enlightenment; and finally,
- when Congress says they have to evaluate.

This may well be because evaluators have not communicated well enough with managers or because the other uses of evaluation have not yet trickled up. These ideas, then, do support the need for more understanding of evaluation among decision-makers, or at least some liaison mechanism, some bridge between evaluators and agency decision-makers.

Third: The Conditions Which Make for Use. I began my exhortation to the panel by listing five conditions for use, with which a good many members of my panel and members of the audience disagreed, but that didn't shake me any. I still believe these conditions are essential conditions.

338

I think in order for somebody to use evaluation results, ·particularly a manager or decision-maker—and this is, I think, all the more true the higher you go in the management hierarchy— that the information to be presented from an evaluation has to be _reliable_. It also has to be _brief_. It has to be _timely_—that is, the information has to be presented at a time whe the manager can use it for a decision. It has to be _comprensible_, no jargon and careful writing. That was, I gather, Chuck Work's prime point last night. It has to be at least to some degree _conclusive_ on some of the questions, if not all of the questions, raised in the first place.

The issue of what kind of structure or organization best promotes the use of evaluation findings gave rise to a great deal of fairly acerbic discussion. One panel member wanted us to stipulate that, for evaluation findings to reach policy-makers, there nee ls to be a centralized evaluation office in the agency, dedicated exclusively to evaluation (i.e., without responsibility for programs) and possessing close and constant access to decision-makers. This was opposed on several grounds:

- that agencies differ in terms of where the power is and where the needs are;
- that people low down in the bureaucratic hierarchy need (and should get) evaluation help, too; and finally,
- that such an organization would ensure only that the basic purpose of evaluation (i.e., public enlightenment) would be foiled because evaluation offices in agencies distort evaluation to suit the purposes of the agencies and the capabilities of the evaluation offices.

It was pointed out also in our panel that no evaluations are ever really conclusive, and that reducing jargon doesn't ensure the

comprehensibility of evaluation findings if decision-makers do not understand evaluation and have not been successfully reached by the evaluators. (Again we came back to the need for a bridge, a mediator between the evaluation and its agency user.) It was also stated that relevance (i.e., again, the "right" questions) and timeliness were more important then conclusiveness. Said one panel member, "Usability is not synonymous with rigor. Some poor evaluations have been used very constructively." Further, there is even some conflict between rigor and use, at least in some cases, because the more an evaluation resembles an experimental design, of course, the more reliable the results will become, but the less likely the evaluation is to be brief and timely and comprehensible.

Our panel did reach some conclusions, and let me state those as I understand them. They weren't shared by everybody, but I think they represent some conclusions by at least a majority of those who stuck with us.

I think those conclusions were that in order for evaluation to be used, the very first criterion is that, before you ever start on an experiment or a program or whatever you want to call it, you have got to find yourself a user, somebody who wants some information. If you don't do that, then you won't ever find anybody who is really going to use it in the end. That seems sort of like a simple proposition, but I think it's one that fairly frequently gets overlooked. As eagerly as users may be sought, they are not often found. Joe Wholey of the Urban Institute said that he has spent a number of years searching Federal agencies for people who considered themselves decision-makers or users and had rarely found any. Since he has done a lot of work for our agency, I assume that reflects on us as well as others.

But beyond not being able to find a user and beyond the necessity of finding one, comes the question of what it is that might be done to improve utilization once you have found a user (if you can find one). That question is a very complex one for which we didn't really find any prescriptive answer which would suit every case, but which involved, among other things, evaluators recognizing that it's essential that they should not merely sit by passively, but seek out opportunities to talk with people in policymaking positions to insist on a role for evaluation, at least insofar as they really believe that that is a proper kind of activity; in short, they should be aggressive about selling their wares. Now, that doesn't mean being aggressive about selling their wares when there is no real need for evaluation, but it does mean that in some respects, evaluators have to understand that in order for anybody to want to have evaluators around and to do evaluations, they have to be useful evaluators which means they have to produce things that people want. Particularly, they have to respond in many cases to short-term questions.

What you have to do, we concluded, is to have a kind of mix; and you have to be able to sell a kind of mix in your agency---a mix of short-term analyses and longer-term inquiries and some assessment and some disciplined judgment and also some evaluations and perhaps some research. That kind of mix is not very satisfactory from the point of view of people who are researchers by training and by inclination. But it is probably an essential kind of activity if the evaluation function is to survive at all as an evaluation function.

We did not, I think, reach a great many other conclusions in particular on which everybody agreed, but we did, I think, conclude that it is essential that there be much greater clarity about what it is that is promised in advance by evaluators about evaluation.

341

There have to be bargains made and negotiations undertaken at many tables. The more complex the Federal program, the more tables to which people must go. Which is to say, if there is a program that involves state and local governments as well as the Federal Government in direct program activity, then there have to be bargains struck all the way through about what the evaluation will do, whom it will serve, what questions it will answer, what it will produce and what kinds of results are expected at what cost. Those kinds of bargains then have to be also taken to another table, which in the case of Federal agencies includes OMB, and also the Congress. There are many bargains to be struck about evaluation. The clearer those can be in advance, the better off the evaluator is likely to be because then he or she will know what it is that it is necessary to produce in the end, and the more probable it will be that evaluations can be relevant, timely, understood and used. Thank you.

END