49587

REPORTS OF THE THREE WORKING PANEL CHAIRPERSONS (CONTINUED)

III. WORKING PANEL I: IMPROVING THE
      INTERFACE BETWEEN AGENCY NEEDS
      AND EVALUATION
      CHAIRMAN: CLIFFORD W. GRAVES,
      Assistant Chief Administrative Officer and Director,
      Office of Management and Budget
      County of San Diego

MS. CHELIMSKY:

Cliff, what conclusions did Panel I come up with?

MR. GRAVES:

Well, the purpose of our panel was to suggest ways of improving
communication between decision-makers and evaluators, on the assump-
tion that such improvement would increase the use of evaluative
information in decision-making.

The panel focused on the evaluator as the most adaptable party:
the decision-maker and the decision-making process were taken as
givens.   Decision-makers make decisions and will continue to do so
whether or not evaluation information is available. Furthermore,
evaluation is only one of several kinds of information that decision-
makers need: political, fiscal, legal and personal information are
other kinds of appropriate input to the decision-making process.
Evaluation supplements, but is not a sutstitute for, these other
forms of evaluation. While this premise was not fully accepted by
all panelists, we agreed that since evaluators appeared to perceive
the problem more acutely, we would have to make the first move.

356

The panel approached its task as a market research problem.
Evaluators have the capacity to supply a product (or service).
This capacity is not infinite, but it is ample:

- The state-of-the-art is highly advanced.
- The Federal government spends substantial funds for evaluation.
- There is a large supply of trained personnel within and available to the government.

(Again, this premise was not fully accepted by all panel members but the disagreement was only a matter of degree.)

Market interest (potential demand) appears to be present.
Increasingly, decision-makers talk as though they would like to have evaluation information; they have supplied increasing resources and status for the evaluation function. However, decision-makers may not fully understand what evaluation is and what it can do.

Decision-maker interest in evaluation is more the result of the growing complexity and openness of the decision-making process and a growing awareness of the shortcomings of other forms of information. In short, decision-makers are interested in evaluation not because they understand what it is, but rather because of the changing environment in which decisions are made.

The panel also accepted the premise that evaluation information will not be used unless the decision-maker wants to use it and that evaluation information should not be used unless the decision-maker knows how to use it.

So, given our ability to supply, and evidence of a market for our product, how should we proceed?

THE APPROACH: FIVE ISSUES

The panel addressed five issues in its search for a "marketing strategy."

1. <u>Who are the decision-makers? Where do they come from and what kind of decision-making environments exist?</u>

   The use of the term "decision-maker" tends to obscure the fact that there are many kinds of decision-makers operating at many points in the process. They vary in the kinds of decisions they can (or are willing to) make, their perspective on any given issue and, therefore, the types of evaluation information they may require. We looked at this issue in two ways.

   First, we focused on the concept of the environment of an agency (or government as a whole) as a decision-making system. The panelists believe that decision-makers are part of a larger system and it is this system, rather than the individual decision-maker, that must be understood. For example, the Congressional Budget Act establishes a system of related decisions and assigns responsibility for those decisions among various elements of the Congress and the Executive Branch. By understanding that Act and the decisions it requires, the key points where evaluation information can be helpful can be readily identified. Similarly, there is a system within each Executive Branch agency.

   The first step in designing an evaluation strategy should be to understand the organization and functioning of the system of interest. Once that is understood, evaluation systems should be designed to fit.

We also discussed individual decision-makers, who come in
all shapes and sizes. We were intrigued by a suggested
distinction made the first day of this symposium, between
"decision-makers" and "position-takers."[41] The latter are
the persons who absorb and analyze information and then
package it for decision-makers. Position-takers are found
in large numbers in the Federal government, and constitute
a good market for evaluation information. They have more
time than decision-makers, and a better understanding of
the analytical side of the evaluation process. They are
conduits for the flow of evaluation information to decision-
makers.

The panel touched on the cascade characteristic of govern-
mental decision-making. At the top are the legislative and
high-level policy-making processes involving relatively few
people and very coarse-grained decisions. These decisions
in turn cascade down through the organization, setting off
administrative decisions. At each level of the cascade, there
is potential demand for evaluation information; different
types are needed, ranging from broad impact and inter-program
effectiveness issues at the upper levels down to operational
efficiency-type questions toward the bottom.

The panelists briefly discussed the importance of under-
standing the incentives that guide the actions of decision-
makers. Much has been written concerning the short-run
outlook of decision-makers, usually in an oversimplified
way. However, it is important when addressing the evaluation

---

[41]
See pages 101-103 above.

needs of particular decision-makers or decision-making levels, that the evaluator understand what makes a decision-maker tick.

Within the panel, a minority view took issue with the panel's approach. That view pictured the evaluator as the seeker of truth, independent of the decision framework, letting chips and decisions fall where they may.

2. What distinguishes evaluation information from other kinds of information needed and/or used by decision-makers? Evaluation information is neither better nor worse than other types of information; it is simply different. The panel spent quite a bit of time trying to determine just what distinguishes evaluation information from other types, and finally identified the following characteristics:

- It is structured information, set within a context, clearly circumscribed. This creates problems of distilling evaluation findings into executive summaries, news articles, and the like; because the first thing to go in such distillations is the context.

- It sets confidence limits, by including cautions to the users.

- It describes and answers questions about a real activity or set of activities according to some theory. It is retrospective, and it addresses specific questions.

- It describes effects against some standards. In fact, often the most important contribution of evaluation projects is the establishment of standards against which activities can be measured.

- It is a feedback loop in a continuing process of
  program development, execution, and refinement. In this
  sense, it is use-oriented.
- Its scope includes consideration of the side effects of
  a given activity, i.e., it is not a closed-end analysis.
- It does not assign values to a given activity, but rather
  tests the activity against values assigned by others.

The panel concluded that the methodology of evaluation is
not its distinguishing characteristic. Evaluation makes
use of many techniques common to other forms of research
and analysis.

3. Evaluation's potential contribution is not fully compre-
    hended by decision-makers; is this because its products are
    badly designed, badly packaged, directed at the wrong
    segment of the market, poorly advertised, or something else?
    As used here, "badly designed" means directed at questions
    of little interest to the decision-maker, or otherwise
    structured to yield irrelevant information. "Badly packaged"
    means that evaluation is not presented in a usable or recog-
    nizable form. "Wrong segment of the market" means that the
    evaluation is not directed at decision-makers or is directed
    at the wrong decision-makers. "Poorly advertised" means
    that the decision-makers are not aware of the existence of
    the information or are unaware of its potential value.

We concluded that the answer was "yes" to all of these. We
then went on to focus on the notion of evaluation as a
threat. The panel believes that the threatening nature of

361

evaluation may be the most important obstacle to its
effective use. Evaluation is a threat because:

- It is public information which, once generated, cannot
  be kept secret or limited to the private use of a
  decision-maker. Thus, it provides persons other than
  the responsible decision-maker with information which
  may adversely affect that decision-maker.
- It is a change force: it seeks ways to improve (change)
  an exis ing set of activities. Change is inherently
  threatei .ng.

To overcome this, the panel believes in the importance of
including "victims" in all phases of evaluation projects
from pre-design and planning through execution and product
packaging. The theory is, the more a project is seen to
be controlled, or primarily usable by the program manager,
the more likely the evaluation information is to be accepted
when it is completed. Also, the more a program manager or
decision-maker knows about how an evaluation project was
put together, the better he is able to implement the
changes recommended. Not incidentally, by giving the
program manager a head start before making evaluation
information public, he is able to accept and perhaps claim
credit for identifying ways to improve his program. As an
example, the Environmental Protection Administration does this
through a task force approach to most evaluation projects.

In the opinion of many of the panel members, evaluation as
practiced at the Federal level is now 99% production and
1% marketing. The lack of attention to marketing (who
needs it and how can it be used?) is a major shortcoming.

362

While apparently common sense, market research is rarely
done before a project is started. This may be because of
poor communication between the evaluation and the decision-
making processes and some uncertainty at the evaluator's
level about intended uses.[42] Even the timing (the point
at which evaluation information may be necessary) is not
always clear. These are obstacles to be overcome, however,
not excuses.

The panel also agreed that the higher up you go in the
decision-making hierarchy, the less time the evaluator has
to present evaluation information and the less the stability
of the decision-making environment. This has two implica-
tions. First, the evaluator should aim at the more stable
elements of the decision-making process, such as Congressional
staff (position-takers) and program managers, rather than
an individual Senator or a Cabinet officer. The second
implication is that in order to secure a significant amount
of the time of the top-level decision-makers, you must get
their attention. This can only be done if they are aware
that a real problem exists. The experience of panel
members was that if the top-level decision-maker is aware
that there is a problem, that decision-maker will take
whatever time is necessary to review information that might
lead to a solution.

---

[42] See Issue 1, pages 358 through 360, also 334 through 336 above.

4. <u>What are the criteria for measuring the effectiveness of evaluation products: technical quality, timeliness, acceptance of recommendations, state-of-the-art advancement, or others?</u>

We've passed the point, as evaluators, where we believe the only standard of a "good" evaluation is whether the recommendations coincide with an actual decision. However, if this is not the principal standard, what other criteria should be used? After all, an evaluation program merits evaluation just as other programs do.

The panel came up with an interesting notion: an evaluation can be considered a success if, according to the evaluator's measures, the program evaluated subsequently improved. The idea here is that, in analyzing the subject program, the evaluator identified or clarified appropriate measures for program performance. If, following the evaluation, the subject program performance improved according to those measures, then the evaluation can be considered a success. (If the evaluation of the program showed that the program was already a total success, then a continued high level of performance against those measures would be acceptable).

To make this assessment requires follow-up to an initial evaluation project, and continuing involvement of the evaluator in the program. This is not usually the case in Federal evaluation programs.

This notion flows out of one of the characteristics of evaluation information noted earlier, that is, it is part of the continuing process of program development including planning, implementation and evaluation.

364

Other factors we identified as important criteria were timeliness and availability (being in the right place at the right time) and whether they raise the consciousness of persons associated with the program to issues of program performance.

5. <u>Given answers to the above, what are the responsibilities of the evaluator, the evaluation manager, the research community and the decision-maker in developing an acceptable product?</u>

There are many players in the evaluation game, each of whom bears some responsibility for an acceptable evaluation product. We kicked around the idea of mandated processes such as the Congressional Budget Act, OMB Circulars, and some of the pending sunset legislation. These have the initial attraction of being action-forcers. However, the panel was not enthusiastic about these as forces to improve the quality and usefulness of evaluation products. Mandated processes guarantee large quantities of evaluation production, but not high quality.

The panel also concluded that evaluators can't do it all, although they can stimulate improvement. The evaluator accepts and conducts assignments but has no institutional responsibility for evaluation planning or for the utilization of evaluation products.

We zeroed in on the evaluation manager--the person or unit responsible for planning, packaging, and disseminating evaluation findings--as the critical factor to the development of an acceptable product. The evaluation manager is the bridge between evaluator and decision-maker. This is the point from which most of the marketing needs to be done; this is the interface.

The panel also noted that the role of the evaluator changes depending on the skills in the agency (assuming that most evaluation is done by outside contractors). Some agencies have a highly sophisticated evaluation process, in which cases the evaluator is more the arms and legs of the agency, carrying out projects designed and pre-marketed within the agency. On the other hand, other agencies lack this sophistication and are, in effect, buying brains, as well as arms and legs. One panel member lamented that the cost per unit of evaluation information should be higher in the latter case, but Federal contracting processes do not recognize the difference.

One of the panel members developed a specific assignment of responsibilities for each of the players in the evaluation game which the panel concluded was a very good one. That report[43] follows directly after this.

CONCLUSIONS

Out of its deliberations, the panel was able to distill its concerns down to just a few points.

The first and most important one is that the approach to evaluation in each agency or decision system must fit that agency or system. There are no universal truths to the design and conduct of evaluation and no universal characteristics to the market. In other words, each evaluation program must be tailor-made.

---

[43]See pages 371 - 373 below.

Second, the most important step toward improving the interface is understanding it. This analysis of the environment and the potential contribution of evaluation is absolutely critical.

Third, evaluation should be seen as part of a continuing loop of program operations, including planning and implementation and evaluation. It is not a separate outside force. Unfortunately, this is not recognized in most agencies.

Having come to these apparently common-sense truths, the panel then concluded that there was little to be gained by further exploring any of them as generalities. Nevertheless, these represent a major agenda for all persons concerned with the effective utilization of evaluation to improve the quality of Federal decision-making.

MS. CHE........:

Do we have some questions? Comments?

PARTICIPANT:

I am Joel Garner, LEAA. I would like as an evaluation manager, or at least as project monitor for evaluation, to reject the idea that I am responsible for bringing coherence to the chaos that we find in terms of the relationship between evaluators and programs. If my office or my personal advancement in the agency is based on that kind of assessment, I need to put out more resumes. I would also like to reject the idea that evaluation is to be assessed on whether the program we are evaluating is in fact successful. Again, if my office or my personal advancement is based on the ability of LEAA to reduce the crime rate (I believe you said that the program itself has to

367

improve after the evaluation), then, if the program doesn't reduce
crime more after the evaluation, the evaluation was not successful.
That's the way I read it. If that's what you said, I think that's a
very dangerous thing to say.

MR. GRAVES:

Let me clarify, but before I do, I can't resist going after your
first assertion. If improving programs through evaluation is not your
responsibility, what the hell _is_ your responsibility?

MR. GARNER:

Well, it's not solely the evaluation manager's responsibility.
There are other people who can be blamed.

MR. GRAVES:

Going onto the second point, what I was trying to get across
(as far as the determination of what an effective evaluation may be is
concerned) is this. If you include utilization as somehow part of
your evaluation criteria, one of the products of evaluation is a set
of measures, perhaps a refinement of measures which already existed.
This is a way of looking at whatever program you are evaluating. That
is really the first thing you do in an evaluation project, and then
you proceed to measure the program's performance against those
standards. You come to certain conclusions about it.

Our view was that if you have done that and you have an effective
evaluation, then the program that you evaluated should somehow perform
better against those measures after you did the evaluation.

Joe Nay sparked this notion with our panel. Perhaps he'd like
to amplify these comments.

MR. NAY:

Yes, in answer to crime, I'd say, well, you shouldn't pick
crime. It's not a good measure. I'd say that, having agreed on a
legitimate set of measures, then you ought to be able to expect to
see improvement in subsequent periods if the program goes on, in
those measures that were used. If crime is a bad measure--and it
is for many programs--then that shouldn't be the measure you are
using. You get that at the beginning--wipe that out at the beginning,
not at the end.

MS. CHELIMSKY:

It seems to me that Blair's point was that there may be more
actors with more roles than Cliff and his panel have suggested.
That is, they are suggesting there is an evaluator, an evaluation
manager and a decision-maker; and there may be a whole lot of other
people who are critical to that process including perhaps somebody
between the evaluation manager and the decision-maker whom we
talked about as being a kind of bridge builder, interpreter,
translator--whatever you want to call him or her.

There are also the people who plan the program and the people
who receive the results of the evaluation, both of whom have some
responsibility for seeing to it that the things that are designed
are things that can be evaluated, at least to some degree. They
have that responsibility. There are people who then have to take
those results and make use of them. So 1 would suggest that there
may be two or three more players in the game.

MR. GRAVES:

I'd agree with Blair. But the point we are trying to make is that, as a matter of fact, I gathered from looking at the roster of people here that most of the people here are evaluation managers or in-house evaluators--somehow responsible for evaluation in-house. There is always a tendency to blame somebody for a problem who isn't around. We had a question early in our panel in terms of why did we have to accept the decision-maker as a given. It's an "all-his-fault" kind of thing.

I made the comment that it's easy to blame the decision-maker because he's not here. But, in my opinion as an ex-Fed, the role and mission of the evaluation division, the Assistant Secretary for Evaluation--whatever it happens to be in an agency--is never clear. I think most of us ended on an optimistic note that maybe one of the things that would come out of this symposium--at least some of the ideas we had--was a clarification, a better understanding on the part of the evaluation manager in terms of what his role and responsibility is.

QUESTION FROM FLOOR:

I would like to back up to Joe's second question. I think that we have taken the assumption that all evaluation is critical, condemnatory. Once in a while on a rare occasion, we find research that is not critical, but that is, in fact, supportive and does not lead to the kind of feedback you are talking about in that we are supporting a homeostatic situation.

MR. GRAVES:

In that case I would say that the criterion would be that it not get any worse, as a result, after the evaluation!

370

ASSIGNMENT OF RESPONSIBILITIES IN THE GAME OF EVALUATION

JOE N. NAY, The Urban Institute
(Member of Working Panel I)

PLAYER: DECISION-MAKER   RESPONSIBILITIES:

- DEVELOP (WITH HIS EVALUATOR) AN UNDER-
  STANDING OF THE DECISION-MAKER'S OWN
  ROLE, NEEDS, AND MEASURES.

    - ACCEPTABLE MEASURES

    - BELIEVED PROGRAM LOGIC

    - MANAGER'S ABILITY, AUTHORITY,
      INTENTIONS TO ACT.

- PARTICIPATE IN CYCLIC CLOSING OF GAPS
  BETWEEN BELIEVED PROGRAM LOGIC AND
  ACTUAL PROGRAM LOGIC.

- TAKE TIME TO UNDERSTAND THE RESULTS.
  (MANAGERS SHOULD NOT EVALUATE THINGS
  WHOSE RESULTS THEY WON'T TAKE TIME
  TO STUDY).

- PARTICIPATE IN DECISIONS ON SEQUENTIAL
  PURCHASE OF INFORMATION.

PLAYER: EVALUATOR

- UNDERSTAND BOTH THE RHETORICAL PROGRAM
  AND THE ACTUAL PROGRAM.

- TIGHTLY RELATE ISSUES, MEASUREMENT
  POINTS AND MEASURES, COMPARISONS, AND
  STRUCTURE OF THE MANAGEMENT AND
  INTERVENTION PROCESS.

- DO ENOUGH PARTICIPANT OBSERVATION TO
  KNOW WHAT IS REALLY HAPPENING.

PLAYER: EVALUATOR   RESPONSIBILITIES:
         (CONTINUED)

- PROVIDE ENOUGH STRUCTURE (FLOW DIAGRAMS?)
  TO SHOW HOW THE MEASUREMENTS TAKEN ARE
  INTERRELATED THROUGH THE ACTUAL PROCESS
  ACTIVITIES.

- CAPTURE EXOGENOUS VARIABLES AND INTERNAL
  FEEDBACK LOOPS.

- BE HEAVILY INVOLVED IN ACTUAL MEASURE-
  MENT.

- CHOOSE APPROPRIATE ANALYTIC TECHNIQUES
  FOR PRODUCING INFORMATION FROM DATA.

- MEASURE, MAKE COMPARISONS, PRODUCE
  INFORMATION, EXPLAIN IT. MAKE RESULTS
  ACCESSIBLE TO VARIOUS LEVELS.

PLAYER: EVALUATION   - HAVE AT LEAST A FEW PEOPLE WHO ARE
         MANAGER (IN   COMPETENT TO DO THE WORK THEMSELVES.
         THE AGENCY)
                            - INVOLVE THEMSELVES IN THE ENTIRE LOOP
                              OF MANAGEMENT, INTERVENTION, AND
                              EVALUATION SO THAT THEY ARE KNOWLEDGE-
                              ABLE IN ALL PARTS OF IT.

                            - FACILITATE AND REQUIRE INTERFACES AT
                              MANAGER/EVALUATOR AND DIRECT INTER-
                              VENTION/EVALUATOR LEVELS.

                            - DO MARKET ANALYSIS AND ASSESS POTENTIAL
                              USERS AND USES, POLICY MARKET, PROGRAM
                              MARKET, INDIVIDUAL MARKET.

                            - DON'T BE AFRAID TO STRUCTURE THE WORK
                              THAT YOU WANT, GET PEOPLE WHO CAN DO
                              IT, AND REQUIRE THEM TO.

372

ASSIGNMENT OF RESPONSIBILITIES IN THE GAME OF EVALUATION
(CONCLUDED)

PLAYER: EVALUATION   RESPONSIBILITIES:
         MANAGER
         (CONTINUED)   • DO:

                                 – EVALUABILITY ASSESSMENTS,

                                 – ISSUES ANALYSIS,

                                 – FIELD WORK,

                                 – SYNTHESIS OF TESTABLE RHETORICAL AND
                                   OPERATING MEASUREMENT MODELS,

                                 – ASSESSMENT OF WHAT IS KNOWN, AND

                                 – DESIGNS AND COSTS FOR KNOWING MORE.

                             • BE THE AGENCY'S CUSTODIAN OF A CONTINUING
                               STORE OF MODELS, KNOWLEDGE, RESOURCES.

CONCLUDING REMARKS

ELEANOR CHELIMSKY,
Department Head, Program Evaluation,
METREK Division of The MITRE Corporation

MS. CHELIMSKY:

Time is fleeting; we're all tired. Let me sum up quickly. I guess in trying to crystallize what I feel has been said at this symposium about evaluation over the last three days, I keep thinking about the citizen reactions to police response time that George Kelling was talking about yesterday. The idea was that if you expect the police to come in five minutes and they get there in ten, you're disappointed. But if you expect them in ten and they get there in five, you're elated.

It may be that much of our dissatisfaction with evaluation today lies not in evaluation, but in ourselves. Many of us find that we are burdened with transactions and activities, that we have less and less time or energy or talent or inclination left at the end of the day either for reflection or for communicating our thoughts adequately. We may be counting on evaluation to fill gaps it was never intended to fill. I think if we expect evaluation to be a surrogate for thinking, as Jim Gregg said, or for problem solving, or for communicating with others, we are going to be disappointed.

If, on the other hand (as Dan Wilner said to me last night in the corridor), we contrast where we are today--in terms of getting acceptance for uses of evaluation--with where we were ten years ago, there is some cause for elation.

It has been said many, many times over the past three days
that evaluation is only a tool, but it does allow something infinitely
precious--the bringing of some rationality into areas which are heavily
charged and counter-charged with emotion and with self-interest.
Clearly we are not going to dissipate all those war-clouds with one
small ray of evaluative enlightenment, nor should we expect to.
We need, as John Evans and Joe Lewis have said, to accumulate evidence
patiently and to help it develop its own momentum.

I think there has been some fruitful airing of divergent, long-
term goals and aspirations for evaluation among us. There has been
perhaps less airing of how to get there from here. We heard Tom Kelly
give some useful definitions of decision-making and position taking,
yet threatened program managers remain a major obstacle for the
integrity of evaluation, for the accumulation of evidence.

We received clear statements from OMB and Congress about their
uncompromising intentions to aggressively pursue the use of evalua-
tion in order to strengthen their review and oversight functions (and
of course, I am thinking here of Joe Nay's first definition of the
term, not his second). What is less clear, however, is how they intend
to do all that; that is, what incentives and sanctions can be, or
will be, used in this area?

The goal of this conference was to confront various points of
view about where we are today in evaluation and to confront them with
candor. I think we have done this, but I am not sure we have done
it completely. Some of the workshops were too big, perhaps, or
another format was needed.

On the other hand, evaluators have not been shy about reproaching
agencies with their managerial sins, and with other sins as well, both of
omission and commission. Agency people have accused evaluators of

gross misdemeanors such as irrelevance, untimeliness, triviality, jargon and over-theologizing, as well as of leading innocent administrators like Chuck Work down the garden path. But evaluators and agency people have also blamed themselves for their own failures. I think the self-deprecating note struck by Jerry Caplan when he articulated the theme of this conference was very helpful. He set the tone for whatever honesty and humility we may have been able to achieve here.

In closing, I'd like to express my appreciation for what I found to be a very open and intellectually stimulating set of statements and interventions by the people here. I know that any conference is the sum of its participants; and if this one has been interesting, it's unquestionably because of the people who have been kind enough to lend us their presence here. Thank you all very much. Cultivate your garden, as Voltaire said, and Bon Voyage to all of you.

END