

Havens

PROGRAM EVALUATION IN SUPPORT  
OF PUBLIC POLICY



094363

BY

Harry S. Havens, Director  
Program Analysis Division  
U.S. General Accounting Office  
Washington, D.C.

G230

At

The International Conference  
On The Future of Public Administration

Quebec, Canada  
May 1979

I am pleased to participate in this conference where we can exchange views on a common problem--how to improve the efficiency and effectiveness of our public programs and institutions. I am particularly pleased to discuss our ideas about the role of program evaluation in support of public policy.

We have already benefited, we hope mutually, from discussions with our Canadian colleagues from the Auditor General's Office and the Treasury Board.

So that you may fully appreciate any biases which are reflected in what I am about to say, you should understand that I am an official of the General Accounting Office, which is headed by the Comptroller General of the United States and

D 46525  
D --- 26  
27

~~710293~~  
094363

which is an independent agency in the Legislative Branch of our Government. I want to underscore this point--that the Comptroller General reports only to the Congress, not to the President--because our experience in a system based on separation of powers may not be fully transferable to a parliamentary system.

#### Need for Better Information

I believe it has become increasingly clear that the need for better information for public policy is not limited to the Executive Branch in our Government, or in others. Legislative bodies have considered a variety of approaches to obtain more relevant and timely information for legislative processes. There are common threads of concern running throughout these various approaches.

- The study of procedures in cost-effectiveness of the Canadian Auditor General reflects the evolution of auditing from accounting issues to questions of "value for money."
- The U.S. Congress is actively considering a Sunset Act which, if enacted, will require the periodic reauthorization of programs based on a careful review of the operation, cost, results, accomplishments, and effectiveness of the program.

- More than half of our State legislatures in the U.S. have enacted some form of sunset laws. These laws all have an evaluation component which is integrated into the sunset cycle.
- Recently, a study by the American Association for the Advancement of Science, reported that "much research and development has already been conducted on measuring governmental effectiveness and productivity, but far more research and development is needed to advance the technologies. Nevertheless, progress in applications need not wait, since the techniques already available have not been used widely."
- About 4 years ago, a group of evaluators from State government in the U.S. formed the Legislative Program Evaluation Section of the National Conference of State Legislatures. One of their first initiatives was to engage the Eagleton Institute of Politics, Rutgers University to develop a clearinghouse for methods of measuring the effectiveness of various State and local programs.

The Need to Involve Program  
Managers in Evaluation for Policy

It is clear that no one has an easy time deciding how to measure program effectiveness. This is difficult even if

program objectives are agreed upon, but if there is no agreement upon objectives, there can hardly be agreement on how to measure effectiveness in accomplishing the objectives. This is not to say that evaluation is useless without agreement on objectives. On the contrary, evaluation of program processes and impacts can be quite helpful to a program manager in improving a particular program operation even though objectives are not clearly stated. However, when evaluation is to be used at higher policy levels and particularly when it becomes involved in political debate and decisionmaking, lack of specificity in objectives and measures can become a very serious problem.

Eleanor Chelimisky stated this dilemma from the program manager's point of view when she said that "many of the agency managers who are to be held accountable are greatly concerned about the methods and processes of their accountability, and about the ability of the evaluation to provide the conclusive determination of program effectiveness upon which the establishment depends. Program managers are disturbed not so much at the idea of accountability, as at their perception of the likely quality of that accountability, the political aspects of that accountability, and the possibility that premature, preliminary, unfair or inconclusive evaluation findings, distorted or even inaccurately reported by the press,

could result in the unwarranted demise of a promising program." This suggests that program managers would be happiest if the evaluation process were designed to serve them, rather than providing a basis for judging them.

There is nothing wrong with program evaluation designed to serve program managers; indeed, the evaluation function should serve their needs, but it is fundamentally wrong to assume that the same evaluations will serve program managers and executive and legislative policymakers equally well. For example, we found in our assessment of evaluation processes in the U.S. Department of Transportation that "in most cases the system is decentralized and evaluative activities, with the exception of a few legislatively mandated studies, are aimed primarily at providing program managers with information on operational and technical deficiencies."

While an evaluation of operational deficiencies will be helpful to a program manager, it is a clearly inadequate base for broad policy review such as would be required by our proposed Sunset Act.

In view of the lack of definition in the past of these different needs for program evaluation, it is not surprising that program managers have not always been enthusiastic about the evaluations of their programs. We believe one solution to this problem lies in developing a more consistent evaluation

policy for Federal programs so that the full range of needs can be met as efficiently as possible.

A Better Basis for  
Realistic Expectations

We are testing concepts for a Federal evaluation policy at the present time in our own assessments of U.S. Government agencies. If the evaluation process is to have a chance of serving the full range of needs, we believe an essential starting point is to determine whether Congress, the executive policy level and the program managers agree on the objectives for the program. We believe that, in our Government, this agreement can come about only through improved communications between responsible legislative committees and the agencies. We have made some specific proposals in a recent report entitled, "Finding Out How Programs Are Working: Suggestions for Congressional Oversight." This report outlines an iterative process during the development of the program through which communication, clarification, and (we hope) agreement should take place.

This contrasts sharply with what has often happened in the past when legislation has created a program and required an evaluation and report, but with no further contact until the report is received, frequently with disappointing results.

More frequent communication would not only provide more assurance that the objectives and effectiveness measures are agreed upon but that policymakers have realistic expectations as to what information will be forthcoming from the evaluation.

Danger of Over Expectations:  
The Black-Box Syndrome

GAO has encouraged and supported various reform initiatives including sunset, zero-base budgeting, and pay incentives for management effectiveness. At the same time, we are concerned about building expectations which cannot be met. We have seen too many promising ideas collapse because of what I have called "the black-box syndrome."

This happens when an interesting and useful idea for improving management is oversold in an attempt to gain sponsorship. The rhetoric takes on an evangelical flavor, claiming that the new technique will somehow solve all problems. The idea gains official acceptance and is implemented. Suddenly, we begin to discover that the new technique does not quite live up to the promises. Disillusionment sets in and the new technique is abandoned because it failed to do things which should have never been expected of it.

U.S. experience with PPBS was a classic example of this pattern. Unfortunately, the same problem continues to plague

efforts to improve public administration. A similar pattern led to the demise of Management by Objectives (MBO) and initial assessments of Zero-Base Budgeting (ZBB) indicate that it, too, is falling short of the initial claims made for it.

A common failing of these efforts was the wholesale installation of a "system" without pilot testing. It deprived us of the opportunity to adjust the concept in the light of real experience. That, of course, reflects the extravagant initial claims, the resulting high expectations, and our normal impatience for results. As weaknesses begin to appear, we continue to be impatient. Rather than correcting the weaknesses, we drop the idea and look for the next bandwagon to jump on. I would like to see us break this pattern, and I believe we must, if we are to achieve genuine, lasting improvements in the way we make and change public policy.

Consequently, I will not offer you any black boxes through which you can solve all your problems. Some of the ideas that we are exploring may warrant wider use. But even if they are successful, a decision to try them elsewhere--particularly in a nation with a different constitutional structure--should be made only after careful consideration. I also urge a period of pilot testing for any ideas you judge relevant, whether ours or anyone else's. Otherwise, you may face something analogous to our experience with PPBS--an idea



which worked well in our Defense Department, but less well when applied mechanically throughout the U.S. Government.

### Congressional Initiatives in Program Evaluation

In fact, however, PPBS in my country is not nearly as dead as it is alleged to be. There is a distinct similarity between the efforts of the U.S. Congress with respect to budget reform, program evaluation and systematic oversight, and what our executive branch was trying to do with PPBS. Efforts in both branches are prompted by a fairly simple truth: society, and government's role in it, are just too complex today for fundamental policy choices to be made in a hit or miss fashion.

To make an informed decision any decisionmaker needs to know:

- What happened in the past?
- Why did it happen?
- What are the options for the future?
- What are the implications of those options?

Ideally, program evaluation will supply answers to those questions. In recent years, the most important stimulus for development of evaluation capability in the U.S. has been the Congress.

In the last 10 years, the Congress has increasingly required evaluation of a program in the laws authorizing that program. In some cases, this takes the form of general requirements; in other cases, the specifications are very detailed. We, in GAO, have encouraged and assisted the Congress in developing such legislative language.

Congress has also acted in a more general fashion. The Congressional Budget Act of 1974 specifically strengthened the legislative evaluation role of the General Accounting Office. It requires us to:

- review and evaluate government programs,
- develop and recommend methods for reviewing and evaluating government programs and activities, and
- assist committees in developing statements of legislative objectives and goals and methods for assessing and reporting actual program performance.

These congressional initiatives appear to reflect a growing recognition of the importance of evaluation as a source of information for legislative decisionmaking, coupled with an awareness that the art of evaluation still has a long way to go.

Guidelines for Making Evaluation  
More Useful to Decisionmakers

Once Congress and the executive have agreed on the

expectations and objectives for programs, successful evaluation still depends on choosing the best methods of data gathering, analysis, and communicating the results effectively. GAO has responded to this need (at least in part) by developing and issuing a series of guidelines intended to make program evaluation more useful for decisionmakers. These guidelines are of several types:

- (1) those for government-wide applicability, such as our suggestions for congressional oversight;
- (2) those outlining ways to analyse broad program or functional areas, such as ways to deal with uncertainty in estimates of costs and benefits from water resource projects; and
- (3) those containing detailed guidance on the analysis of specific programs, such as ways to compare the cost and performance of our Section 236 rental housing programs and other Federal housing programs.

I believe there is increasing recognition of the need for such guidelines if the evaluation process--and its professional practitioners--are to command sufficient credibility to be relevant to public policy formulation. For example, the recently established Evaluation Research Society of America took as one of its first priorities the development of evaluation

standards and created a Committee on Evaluation Standards for the purpose. We welcome these efforts and GAO staff have participated actively in the Evaluation Standards Committee.

Remembering full well that I promised not to propound any "black boxes," I would only suggest that you consider the relevance of some sort of standards by which to judge the quality of analysis in your own environment. In considering this matter, however, I caution you to think about the trade-offs which are entailed. If the standards are too rigid, you may stifle innovation in analysis. If they are too rigorous, you may find that no one who is working in the real world can meet those standards. On the other hand, if the standards are too loose (or if there are no standards) you will find some pretty shoddy work passing as analysis and being used to make important policy decisions.

There are undoubtedly countries represented at this conference where it would be inappropriate for a government agency to have the authority to issue guidelines for evaluation which, in the U.S., is vested in the Comptroller General. For those countries which want a standard-setting process, but are reluctant to see that authority vested in a government agency, the Evaluation Research Society may provide a model to be explored. It is a model which is found in many other

professions. For example, in many (perhaps most) countries, the accounting profession largely sets its own professional standards.

Serving the Evaluation Needs of  
All Management and Policy Levels

I have alluded earlier to the concern of program managers about being held accountable through the evaluation process, given the current state-of-the-art of evaluation. I am hopeful that our guidance will begin to alleviate this concern. In addition, however, I believe the confidence of program managers will be improved by current developments in the U.S. Office of Management and Budget. OMB has recently issued a new directive on Management Improvement and the Use of Evaluation in the Executive Branch. The circular provides guidance on management improvement initiatives designed to increase the efficiency and effectiveness of program operations in the executive branch. It also emphasizes the importance of evaluation both for overall management improvement efforts and for the budget process.

Perhaps the most important area of agreement between OMB and GAO is covered in the following statement from the OMB directive: "the heads of all executive departments and agencies are responsible for developing and pursuing comprehensive management improvement efforts. The objective of such efforts

should be discernable improvement in Federal programs--in the efficiency of administration or management and in the effectiveness of results. The basis for identifying management improvement needs is a sound evaluation system." This is fully consistent with our own views.

I am particularly pleased that OMB recognizes the need to review agency evaluation activities themselves. The directive requires these reviews to:

- weed out evaluation efforts that do not contribute directly to improving programs;
- assure a balanced emphasis on both evaluation and prospective analysis, such as planning and policy analysis; and
- assure that available evaluation resources are used economically and efficiently.

These are very similar to the criteria we use in our own reviews of agency evaluation systems. We have issued reports on the evaluation systems in the two major departments and have a third in process. As a result of this work, together with a broader assessment of evaluation policy in eight major Federal agencies, we plan to develop and issue a guideline on evaluation management and policy. I expect this guidance to be built on two premises: (1) program evaluation is a fundamental part of effective program administration,

and (2) the best test of the effectiveness of the evaluation function is improvement in the performance of the programs being evaluated. It is clear that evaluation, as a function, involves more than just those activities which include the word "evaluation" in their title. Those who decide what to do with the results of an evaluation--the program managers and policymakers--are also part of the evaluation process.

### Conclusion

The governments of the nations represented at this conference differ in many fundamental ways, but we share a common need to make our public institutions more effective and efficient in meeting the needs of our citizens. There is an endless list of the things we must do to accomplish that goal, but somewhere on that list must appear the ability to understand better the results we are achieving with public funds and to find ways of improving those results. This is another way to describing program evaluation and we must be about the task.

I do not claim that we know how to do it as well as we would like; our list of questions is much longer than our list of answers. I hope that by sharing our experiences, we will find that each of us will be able to answer questions which have eluded the others. There are enough unanswered questions to keep us all busy for many years to come.