The Congress and Evaluation*
Prepared for Delivery by
Harry S. Havens
Assistant Comptroller General
of the United States
at the
Sixth Annual Research Conference
of the
Association for Public Policy Analysis and Management
New Orleans, Louisiana
October 19, 1984

INTRODUCTION

Congress is a collegial, deliberative, decisionmaking body. Evaluation is a function, one of whose principal values lies in its use as an aid to the decisionmaking process. If one accepts both of these statements, it is difficult to escape the notion that Congress ought to be a major user of evaluation. Yet the scholarly literature on the Congress is dominated by carefully documented studies, the conclusions of which all seem to demonstrate that congressmen's votes are determined by factors unrelated to rigorous analysis of costs, benefits, or effectiveness.

One school of thought holds that voting behavior is dominated by the ambitions of the incumbent. Legislators vote in line with the interests of the constituency they hope to retain or are seeking to capture[1]. Another theory is that members make their decisions by comparing the policy content of an issue to a set of predetermined policy positions. These policy positions, it is argued, are quite stable over time. As a result, Congress changes policy by changing members, not by members changing their minds[2].

How can it be that a decisionmaking body can ignore what is arguably, at least, a highly potent tool of the decision process? This, too, has been examined on several occasions. One view is that legislators, who represent and are concerned about individual constituencies, simply have little interest in the societal costs and benefits which tend to be the focus of what is often considered to be the mainstream work of program evaluation and policy analysis. Haveman, for example, asserts that "Policy analysis answers questions that legislative policymakers are not interested in either asking or having asked"[3]. Another view is that members do not gather much of their information from written products, which are the predominant communications modes in evaluation. Instead, they collect information, largely orally, from staffs, other members, constituencies, interest groups and the administration[4].
This base of analytical work could lead one to be rather pessimistic about the likely impact of evaluation in a congressional setting. And yet, somehow, the notion persists that Congress, a decisionmaking body, should find some use for a decisionmaking tool. Indeed, that notion persists in the Congress itself. It appears in statutory form in Title VII of the Congressional Budget Act of 1974[5]. Nor was this a unique event. Over the past decade or so, there have been numerous requirements levied by Congress (some by statute, some by committee report language) for agencies (including both operating agencies and independent agencies such as the General Accounting Office) to perform evaluations and report the results to Congress. In another context, the various proposals for "sunset" legislation have usually contained some provision for systematic evaluation of most government programs.

It is difficult to dismiss all of these actions as meaningless. No doubt some of them represent rhetorical gestures. But some cases, at least, seem to reflect a serious desire for useful evaluation. And it is possible to reconcile such a desire with existing data on congressional voting behavior.

Not all congressional decisions are of the sort to be governed by the influences of ambition or policy predisposition. And not all evaluation effort is devoted to the major political questions on which these influences are likely to prove most
powerful. The term "evaluation", of course, encompasses a wide range of work, from relatively simple assessments of process to very elaborate and complex analyses of program impact. Despite the broad range of possible evaluations, however, most of the work of evaluators deals with the question of how to make a program work better. Such evaluations presume the legitimacy of objectives, rather than challenging them. The recommendations, therefore, focus on improvements at the margin. Only rarely does an evaluation yield evidence which is sufficiently reliable and conclusive to challenge (or validate) the very existence of a program. Most evaluations, therefore, deal with issues which, in principle at least, should not require legislators to rely on political imperatives and ignore the results of the analysis.

In this context, even Haveman's argument that legislative policymakers are not interested in societal costs and benefits loses some of its force. In speaking to marginal improvements rather than global costs and benefits, evaluation is a tool for reducing the cost of a given stream of benefits or for achieving greater benefits at the same cost. Given the present fiscal environment, this is an area in which members are likely to be interested, regardless of their political views of the program as a whole.
Why, then, does the scholarly literature seem to allow so little room for the use of evaluation and analysis by the Congress? And why do evaluators--at least many of them--still bemoan the failure of Congress to make effective use of their work? There are several answers. One is that more use is being made of the work than is commonly recognized. Another is that evaluators' goals with respect to the nature and extent of utilization are unrealistic. And still another is that many evaluators have not yet learned how to communicate their work in such a way as to encourage use. Each answer contains an important element of truth. Understanding those answers requires some examination of how Congress actually uses evaluation.

THE INSTITUTIONAL CONTEXT

Congress, as an institution, has a unique set of relationships with the process of program evaluation. It is a collegial, deliberative body whose processes depend on the ability to collect, process and absorb a wide variety of information from a wide variety of sources, to reconcile or compromise diverse points of view, and to reach decisions in the form of enacted statutes.

Congress, however, is not just an institution with institutional behaviors. Congress is also a collection of 535 individual voting members. As discussed above, these members,
and their staffs, are guided by a wide variety of motivations, relating to personal values, principles and ambitions.

The literature which seeks to analyze this set of behaviors—and the motivations underlying them—is dominated by examination of voting records. These are primarily votes cast on the floor although some attention has been given to recorded votes in committee[6]. This is hardly surprising, since voting records represent a readily available body of data which can be subjected to various forms of useful statistical analysis. At the same time, however, this focus tends to ignore some significant activity which may not be reflected in recorded votes. Among these is the frequently informal committee and subcommittee markup process.

Experience in the General Accounting Office has been that direct, instrumental use (the adoption of one's evaluative recommendation) is much more likely to occur in committee deliberations than in the context of a floor debate and vote. Decisions made at the earlier stage, particularly those involving marginal change in the design or operations of a program, may be taken in a consensus fashion, without a formal vote. These decisions, therefore, would not be reflected in the statistical data which is available for analysis. There is some evidence that those who participate actively in committee deliberations
occasionally change their views, even on major issues[7]. It is reasonable to expect somewhat greater susceptibility to argument and analysis on issues to which a member attaches only minimal political significance.

The author's experience has been that most legislators distinguish reasonably clearly between decisions which have a conspicuously political content and those which do not. On decisions which are recognized as involving important political considerations, and on which action must be taken formally and in public, behavior seems quite consistent with the literature. Behavior on other issues is frequently quite different. On these issues, a member is often open to being convinced of the "best" solution on the basis of analysis.

The same distinction may manifest itself in other ways. In developing a new program, for example, the congressman may have relatively unshakable views on the general shape and direction of the program. At the same time, others (the staff, experts, et al.) may be left to fill in the details. Political significance is seen as being attached to the broad dimensions of the program, not to its details[8].

Decisions in committee on issues involving details of program design and management are frequently accepted without debate or vote at later stages of the legislative process[9].
The evidence is that other members consciously defer on these details to those who participated in the committee hearings and markup process. Thus, the decision process on the type of issue which is addressed in most evaluations is hidden from the statistical analysis underlying most recent work on congressional decisionmaking.

**FORMS OF USE: INSTRUMENTAL**

Given this context, what strategies should the evaluator pursue if utilization is an important objective? At this point it is helpful to distinguish various forms of use[10]. Much of the time, an evaluation is considered to have been used if its recommendations are implemented, representing instrumental use. This is certainly one form of use. Often it is the only form which is relevant. But other forms are possible and, for some types of evaluation and in some circumstances these may be more significant. One such form is persuasive use, in which recipients use the evaluation as evidence with which to convince others of the correctness of a particular position. Another form is conceptual use, in which recipients make the content of the evaluation part of the general intellectual framework with which they approach an issue or set of related issues.

A member's dichotomous decision framework has important implications about expectations for use in a legislative setting.
It suggests particularly that these expectations should take into account the political significance of the evaluative conclusions and recommendations. The greater the extent to which the evaluation deals with issues which a member will see as involving political considerations, the less likely it is that the evaluation will be used in an instrumental fashion and the more likely it is that the relevant form of use is persuasive or conceptual. Instrumental use of an evaluation on politically significant issues is certainly possible, but analysis of congressional decision processes suggests that it would be relatively unusual. Evaluators should be realistic in this regard and not be unduly disappointed if it fails to occur. On the contrary, they should anticipate that situation and look for strategies that will facilitate persuasive or conceptual use.

At the same time, the possibility of instrumental use on politically sensitive issues should not be dismissed out of hand. Issues may be politically sensitive to some members and not to others. And some members may have more political flexibility than others. Thus, on many issues, there will be a group of legislators who are open to persuasion. If that group is large enough to hold the balance of power, a persuasive piece of analysis can affect the outcome. While such an affect is unusual, examples can be found.
Assuming that the evaluation involves a question on which instrumental use can be reasonably expected, there are strategies which an evaluator can pursue to encourage such use. One is to choose a topic and a focus which is known to be of interest to one or more influential members. An evaluation of this sort may be initiated by either party, but often culminates in a relatively explicit agreement between the user and the evaluator on the scope of work and timing and form of the resulting product.

Another approach, one which can be combined with the first, is to work closely with committee staff in both the planning and implementation stages to assure that the work will be responsive to their needs and that they understand the content and implications of the evaluation. There are two important reasons for establishing this relationship. On issues which are not politically sensitive, staff frequently have substantial authority to make decisions. In principle, such decisions are no more than tentative, subject to approval by the member. In fact, however, they are often accepted without substantial discussion[11]. Another reason for working through the staff is the relative accessibility of these individuals, and their influence with the member. Even on politically sensitive issues, a highly respected staff member may convince a legislator to alter his/her position if the staffer is armed with a convincing analysis.
Whether working with staff or with the legislator, instrumental use is much more likely if the evaluator is attentive to key elements in the evaluator-customer relationship. These include reaching agreement on a clearly-defined, relevant, answerable question; maintaining contact with the customer during the study; and searching for the most effective way to communicate the results[12].

PERSUASIVE USE

Although instrumental use is the goal to which most evaluators aspire, it should not obscure the significance of use in a persuasive or conceptual fashion. Both forms of use exist, although the evidence is both ambiguous and elusive. Persuasive use may actually be difficult to distinguish from instrumental use in a particular case. Suppose, for example, a member cites an evaluation report as the reason for a particular action. That is certainly evidence of use. But it would be misleading to assume that such a statement is strong evidence of instrumental use. The statement may well be an example of persuasive use, citing credible authority in support of a decision the member would have made even in the absence of evaluation.

If one is only searching for evidence that evaluations are used, the form of use may not matter. But if one is seeking
to maximize a particular form of use, or to discern the consequences of an evaluation with some precision, the distinctions become significant. From an analytical perspective, the distinctions can probably be made most reliably by reference to the content of the decision. The stability of congressional voting behavior on major political issues becomes an essential assumption in such an analysis. If an evaluation is cited in connection with an issue involving substantial political controversy, it is relatively unlikely that instrumental use is involved. Unless the legislator is voting in a fashion which is a departure from past voting patterns, the citation of an evaluation can be reasonably presumed to represent an example of persuasive use.

It is important not to assume that persuasive use is improper or somehow demonstrates that bias was present in the evaluation. A perfectly sound evaluation may be used in persuasive fashion. If the evaluation is sound and the use does not involve distortion of the results, persuasive use should not reflect adversely on either the evaluator or the user.

Another common example of persuasive use is seen when a report is used by a legislator to help explain an action to
others. Kingdon, for example, points out that "Congressmen are constantly called upon to explain to constituents why they voted as they did"[13]. There is the same need to explain or justify actions to other members and to political or administrative officials in the executive branch. In carrying out this function, a legislator might either use an evaluation report as a source of information or simply cite it as support for the action. In this case, the member's views may not have been changed by the evaluation, but the evaluation may have made it easier to act on those views by making it easier to explain or rationalize them.

If an evaluator wishes to facilitate persuasive use, effective communications should be a central element of the strategy. The objective should be to make it as easy as possible for the user to understand and relay the primary message of the report. This means capturing that message accurately in a few words or phrases through the use of descriptive titles and highly condensed digests or executive summaries.

CONCEPTUAL USE

Finally, there is the case of conceptual use. If an evaluation caused a member to see a significant, but previously unrecognized aspect of an issue, it would be a clear case of conceptual use. While probably unusual, such an event would be consistent with the literature on congressional voting behavior.
In the policy dimension framework, for example, an evaluation might reveal a policy dimension which was previously hidden. Alternatively, an evaluation might reveal a previously unrecognized effect on the constituency of interest. Admittedly, there is little evidence of such direct conceptual use. In part, of course, the lack of evidence reflects the inherently elusive nature of conceptual use as a target of analysis. The phenomenon of interest occurs in the mind of an individual, where it cannot be directly measured, and reliable surrogates are difficult to identify.

It seems likely, however, that the conceptual use of evaluation commonly occurs in an indirect fashion. The results may become part of the flow of ideas from the academic community, the business community or the media. In these forums, of course, an evaluation competes with other work having similar or conflicting messages and the eventual outcome is likely to represent an amalgam of these ideas.

While conceptual use is extremely difficult to demonstrate or measure in most cases, it clearly exists and can be very powerful. The history of economic policy-making in this country provides excellent examples. Keynes' work supporting the concept of countercyclical fiscal policy was first published in the mid-1930's[14]. Those ideas quickly became the center of debate among theoretical economists.
That debate, over the ensuing years, yielded a plethora of books and articles interpreting, testing and amplifying the key elements of Keynesian theory. By the 1950's, the cumulative effect was such that Keynesian economics dominated the teaching of economics and the formulation of economic policy[15]. Today, Keynes' ideas (at least the most important ones) are so thoroughly imbedded that we no longer recognize their sources. The debate over the federal budget deficit, for example, is overwhelmingly Keynesian in its context.

Needless to say, Keynesianism is not the only factor in economic policy-making. Keynes' tendency to downplay the role of monetary policy (a tendency which was carried much farther by some of his more extreme followers) made room for an alternative view. The monetarist view, expressed with particular effectiveness by Friedman[16], has had a history which is similar in many ways to that of Keynesianism. The basic theory was stated, widely debated among theoretical economists and then began to appear as a basic for policy. With Keynesianism, the policy instrument was the budget. With monetarism, it was the decisions of the Open Market Committee of the Federal Reserve. In each case, however, conceptual use ultimately was reflected in the intellectual framework and decision processes of the Congress. For Keynesianism, that reflection was the Congressional
Budget Act of 1974, with its focus on the budget aggregates. For monetarism, it was the requirement that the Chairman of the Federal Reserve Board appear regularly to report on the behavior of the monetary aggregates[17].

As numerous observers have noted, congressional control over monetary policy is much less complete and direct than in the area of fiscal policy[18]. In large part, this is attributable to the firm tradition of an independent Federal Reserve and the relatively recent general recognition of the importance of monetary policy in managing the economy.

To some of us, at least, the next evolution would seem to be a successful blending of the two theoretical frameworks. (To my way of thinking, the central theories are quite reconcilable. Some future Nobel laureate is probably developing that reconciliation as we speak.) That blend, after an appropriate period of gestation and acceptance in theoretical circles, will eventually find its way into policy processes and be reflected in greater integration of monetary and fiscal policies.

Few of us, if any, can aspire to the sort of conceptual use which was made of the work of Keynes, Friedman and their followers. Most of us spend most of our time working at a somewhat more pedestrian level. Nevertheless, we should not
lose sight of the fact that our work often becomes part of the intellectual environment and may have its most significant effects, in the final analysis, by altering that environment.

The indirect nature of much conceptual use suggests strategies for facilitating it. The central objective should be to assure that those with the power to influence thinking are made aware of the work. There is, of course, a wide variety of possible dissemination strategies. The choice of which to pursue should depend on who is to be reached. Each potential audience represents a possible channel back to the Congress. But each audience has its own standards of credibility and preferences on form. Failure to respect such standards and preferences entails the risk (indeed, the likelihood) of failure to reach and convince that audience.

To maximize conceptual use, it may prove necessary to pursue several dissemination strategies simultaneously, each carefully tailored to the needs of one or more target audiences. Some may like frequent (brief) progress reports on work in process, while another group may be satisfied only by a detailed, technically complete final report and a third may not look at anything but a highly condensed executive summary. This sort of communications strategy may seem rather expensive, but the expense may well be unavoidable if major conceptual use is to
occur. In addition, of course, the expense of a well-developed dissemination strategy is generally quite small compared to the cost of performing the evaluation itself.

**IMPLICATIONS FOR EVALUATION**

If the general assessment of use set forth above is correct, it suggests (among other things) that the types of evaluations in which Congress is interested may differ significantly from those which the evaluation community as a whole has tended to regard most highly. It helps explain why, with some notable exceptions, Congress has not yet displayed much enthusiasm for sophisticated, methodologically rigorous impact evaluations. On the other hand, the use of evaluation for persuasive or explanatory purposes is consistent with the fact that members display substantial interest in assessments of the extent to which program services are delivered to the intended target groups. Similarly, the absence of overriding political imperatives when considering incremental changes to program design and operations helps explain the continuing interest in process-oriented evaluations designed to yield management improvements and greater administrative efficiency. The strong, continuing interest in evaluations dealing with service delivery and process efficiency is quite evident in GAO's relations with the Congress. Such evaluations are often requested and, when
produced (either by request or under GAO's general charter), result in an agreeably high frequency of instrumental or persuasive use.

It is unfortunate that evaluators have paid so little attention to the decision processes of their legislative clients. Indeed, some evaluators would undoubtedly consider serious curiosity on the subject to be inappropriate. One of the results, however, has been that the activities which build stature among evaluators are often not those which are most useful, at least in a congressional environment. Evaluations aimed at improved management efficiency and those which seek to identify the extent of service to target groups do not yet garner a great deal of respect in the evaluation community. Evaluators have been encouraged by some of their peers to look down on anything short of a large scale controlled experiment dealing with program impact. Yet it is precisely this sort of evaluation which is most likely to encounter the political imperatives which militate against the instrumental use of the results. The controlled experiment is powerful as a basis for scientific inquiry; in the congressional arena it seems substantially less powerful.

This is not intended to demean the large-scale controlled experiment, or to suggest that it is irrelevant in the congressional context. The Congress itself has participated actively in
decisions to engage in such experiments. The Income Main-
tenance and Housing Allowance experiments are obvious examples. But one's expectations about the likely impact of such efforts should be realistic. It is not obvious that either of these major experiments has yet had much effect on the nature or design of the programs to which they are relevant. (It is too early to tell if the Administration's housing voucher proposal will lead to substantial changes based on the results of these experiments.)

Some may attribute the relatively modest policy effect of large scale social experiments to inadequacies in their design or execution. Flaws of this sort have certainly been identified, as well as other limitations on the extrapolation of experimental results to the construction of social programs[19]. Experience in recent years seems to suggest that these problems may well be almost inescapable. Acknowledging that we are unlikely to see the "perfect" large scale social experiment lends some credibility to the view that imperfections in the research have caused the loss of policy effect. In the author's view, however, the nature of decisionmaking processes explains the absence of policy effect in a much more direct fashion.

The basic structure of income transfer programs--which was the issue to which these experiments were addressed--is a profoundly (and properly) political issue. In that environment,
it would be unreasonable to expect any single analytical effort to have a major effect in the short run. As Schick suggests, "Everything is grist for the congressional mill, and analysis enjoys no preferred position by virtue of its esteem in intellectual circles"[20].

The most that should realistically be expected is that the knowledge gained from these experiments will become part of the intellectual framework of the country over the long run. If this sort of conceptual use occurs, the ultimate effects of the work may well be of great importance. It is unlikely, however, that the cause and effect relationship will ever be more than speculative.

There are other factors, apart from political considerations, which limit the short-run, instrumental use of large-scale impact evaluations. These projects yield complex reports full of ambiguities. Making use of them often requires quite a high level of understanding of the program. Interpreting them properly may well require that the reader also have a rather sophisticated understanding of the research process. This, in turn, requires an investment of time which few legislators can afford. Impact evaluations also tend to take a long time to complete, longer than most congressmen are usually willing to wait for an answer.
There is, and will continue to be some interest in impact evaluation. That interest may well stem, in large part, from the presence of technically trained committee staff, who have the time and interest to follow complex evaluations. There seems to be a trend to increasing the use of such staff. But the constraints on instrumental use are rather severe. As a result, large-scale impact evaluation seems likely to remain less in vogue in a congressional setting than elsewhere in the evaluation community.

Evaluators are correct in observing that much evaluation activity goes unused in the Congress. Many of the reasons lie in the nature of the Congress, which governs what its members perceive as useful. If evaluators wish to see greater use made of their work, they would do well to spend more time studying the way their various clients (legislators, staffers, interest groups, et al.) function and interact, and then produce the sort of evaluation which is appropriate to that context.

Evaluators would also be well advised to think about the appropriate form of communication for the congressional context. That they have not yet done so is most evident in the continuing emphasis on written communications. The evidence in the literature is quite clear that most legislators receive very little of their decision-related information in writing[21]. This view would be confirmed by experienced observers who would
point out that this is not necessarily a matter of choice. Most congressmen simply do not have time to read about any subject at length, a characteristic which is equally prevalent, by the way, among decision-makers in the executive branch.

It is possible, however, to develop somewhat more effective means of communicating. Experience suggests that oral communications, for example, are more effective than written, particularly in a congressional context. Experience also suggests that informal briefings, with opportunity for substantial give-and-take, are more effective than formal hearings. For the evaluator, these techniques may appear less efficient than the written report. An informal briefing can reach only a small number of potential users, while a written report can reach them all. But a report which goes unread has reached no one, while a tailored briefing may reach a few influential members quite effectively, and through them, a much broader group.

In summary, if evaluators give more attention to the nature of the client, they will realize that their work is used more extensively and in a more diverse fashion than they may realize. They will also find ways to increase that use, by doing work of the sort the client is most likely to use and by communicating it in a more effective fashion.
REFERENCES


5. Public Law 93-344, enacted July 12, 1974. Title VII carries the label "Program Review and Evaluation." (88 STAT. 325)


7. See Kingdon, op. cit., especially the discussion on p. 82 of the reasons for reliance on fellow congressmen with expertise.

8. The boundaries of the "political" can seem odd. What came to be known as the Model Cities Program, for example, was originally called the Demonstration Cities Program. The name was changed, at least in part, out of concern for the possible repercussions of using the work "demonstration" in the political environment of the mid-1960's.

9. Woodrow Wilson expressed this view particularly strongly when he observed that Congress usually meets "to sanction the conclusions of its committees as rapidly as possible." Wilson, Woodrow, Congressional Government. Houghton Mifflin Co., New York, 1913, p. 78.

11. There is substantial dispute over the importance of staff in the decisionmaking processes of Congress. Kingdon's data suggests a relatively minor role. On the other hand, Malbin, Michael J., in Unelected Representatives, Basic Books, New York, 1980, indicates a much more influential role on the part of the entrepreneurial, "new style" staff. If Malbin's thesis is correct, it has significant implications for communicating evaluation results usefully. The staff is invariably more accessible than the member.


13. Kingdon, op. cit., p. 46. Later in the same section, he discussed strategies which congressmen use in the process of explaining. One of these strategies involves citing respected authority. While not stated, such authority might include a credible evaluation report.


17. Public Law 95-188.


