An area of concern for evaluation research is the extent to which evaluations are utilized in policy and program development and change. The current review critically discusses definition of utilization and methods for detecting utilization. Five "clusters" of variables have been found to affect utilization. These are: (1) the relevance of evaluation to the needs of potential users; (2) extent of communication between potential users and producers of evaluations; (3) translation of evaluations into their implications for policy and programs; (4) credibility or trust placed in evaluations; and (5) commitment or advocacy by individual users.

RESEARCH ON THE UTILIZATION OF EVALUATIONS

A Review and Synthesis

LAURA C. LEVITON
University of Pittsburgh

EDWARD F.X. HUGHES
Northwestern University

*A*n issue of concern in evaluation research is the extent to which findings of studies are utilized (Young and Comptois, 1979). The rationale for conducting evaluations has frequently been their usefulness in informing policy or improving programs (Weiss, 1973). The claims for usefulness of evaluation are undermined, however, when results are ignored by decision makers (Cook, 1978). Concern over apparent nonutilization of research information extends beyond evaluations to all of policy-relevant social science (Lynn, 1978). This article reviews information about the use of evaluations for programs and

policy. The review consists of three parts: (1) a critical discussion of definitions of utilization, (2) a discussion of methodology, and (3) a review of variables that have been found to affect utilization.

There is a dearth of studies dealing with the utilization of evaluations per se. This review focuses on evaluations, but draws on the larger body of information about the use of social science, in general (Caplan et al., 1975). This information provides potent insights into the use of evaluations, but inferences should be drawn from it with caution. Evaluations differ from other social research used in government by being more often politically sensitive (Weiss, 1973; Campbell, 1969). Few researchers have adequately distinguished the utilization of evaluations and other forms of research (but see Weiss, 1978; Young and Comptois, 1979). A second source of information is the writing of experienced evaluators and policy makers about the nature of utilization and ways to enhance it. These suggestions corroborate research findings and suggest areas for further research but are limited by the experts' orientation and bias. Writers on the subject of utilization have also long recognized the relevance of other traditions in political science and organization behavior.

If findings from several of these sources corroborate each other, and if a range of methodologies are used, one can begin to assess the convergent validity of conceptions of utilization and variables that affect it. For example, the importance of user involvement in research is supported by empirical data (Patton et al., 1977; Windle and Bates, 1974), expert opinion (Chelimsky, 1977), and research on organizational change (Glaser, 1976). This convergence allows us some tentative generalizations.

DEFINITION OF UTILIZATION

Utilization is confined here to use of evaluation results for programs and policy only, not by academicians or by the press, for example. Nor is the related concept of misutilization discussed; the reader is referred to papers by Cook and Pollard (1977), Cook et al. (1980), and Mushkin (1973). However, the current review points to situations that pose a danger of misutilization.

"BOTTOM-LINE" CRITERIA FOR UTILIZATION

We believe there are two necessary criteria for all types of utilization described below. The presence of these criteria must be inferred from observables, just as the hypothetical construct, "utilization," must be inferred. First, Cook and Pollard (1977: 161) suggest that, to consider evaluations used, there must be "serious discussion of the results in debates about a particular policy or program." That is, there must be an attempt to relate the findings to the policy or program issue under study or evidence that the results have been translated into their implications for these issues. This process is what Weiss (1978) and Weinberg (1979) term "information processing." The reading of evaluation reports, for example, is not utilization, although Lyon et al. (1978) report that some officials consider it utilization.

We would add a second criterion that is slightly more exclusive than the one advanced by Cook and Pollard. To be considered used, there must be evidence that in the absence of the research information, those engaged in policy or program activities would have thought or acted differently. After all, people could give serious consideration to information and then deliberately choose not to use it. Measures of this criterion could include changes in certainty about a decision or belief (the criterion used by Patton et al., 1977), changes in attitudes on the relevant issue, increase in the forcefulness of an argument, and action directed at changing or preserving programs and policies, that take their form in part from evaluation evidence. This may be difficult to verify, but would confirm our belief that utilization had occurred.

UTILIZATION DIFFERENTIATED FROM RELATED CONCEPTS

These criteria help to distinguish utilization from *impact* on programs and policy, on the one hand, and *utility* of evaluations, on the other. Impact may be defined as modification of policy or programs, to which evaluation findings have contributed. The potential utility of an evaluation involves the relevance of the findings to issues of concern in policy and programs (Huron Institute, 1979; Lynn, 1978; Schmidt et al., 1977). Most writers now agree that use is different from impact (Cook and Pollard, 1977). Investigators have found relatively modest impact of social research, but many instances in which people attempted

to apply research findings to policy or program issues (Caplan et al., 1975; Patton et al., 1977). Few have confused utility with utilization—in fact, studies of utilization have been motivated by the nonuse of results with high utility (Fairweather et al., 1974).

Barriers to the eventual achievement of impact relate to each of these three concepts. Evaluation may have low utility because it may be of low quality or its implications may not be practical. An evaluation may not be utilized because its utility is not recognized, or people may choose not to use it. Finally, an evaluation may have no impact, because impact is dependent on decisions by many individuals in government who may fail to see utility or choose not to utilize evaluations.

CATEGORIES OF UTILIZATION

So far, we have discussed utilization without reference to the purposes to which it is put. There is a growing feeling that utilization is not a unitary concept, that there are different types of utilization, based on the purposes they serve. A more finely grained set of definitions that meet our "bottom-line" criteria are probably more informative than a blanket definition (Cook, personal communication, 1979; Larsen and Werner, 1981). Researchers have identified three broad categories of utilization that can be distinguished by their purposes: instrumental, conceptual, and persuasive (the last category is termed "symbolic" by Pelz, 1978; and Young and Comptois, 1979).

Instrumental use was defined by Rich (1977: 200) as "cases where respondents cited and could document . . . the specific way in which research was being used for decision-making or problem-solving purposes." Examples of instrumental use of evaluations in government are decisions reached jointly by the Department of Education and National Institute of Education to fund or not fund dissemination of educational innovations, based on evaluations of their effectiveness (Tallmadge, 1977). Conceptual use of research information was defined by Rich as "influencing a policy maker's thinking about an issue without putting information to any specific, documentable use" (1977: 200). An example of conceptual use of an evaluation in the health field is revision in thinking about the ability of Professional Standards Review Organizations to contain hospital costs, unless other cost-containment measures are instituted as well (Dobson et al., 1978).

The third category, here called persuasive use, involves drawing on evaluation evidence in attempts to convince others to support a political position, or to defend such a position from attack. Examples can be

found in congressional staffers' use of evaluation evidence in building a coalition in support of legislation (Florio et al., 1979). The difference between persuasive use and the other two categories listed above is that persuasive use involves interpersonal influence, getting others to go along with the implications of evaluation.

EMPIRICAL SUPPORT FOR THE CATEGORIES

Rich (1977) studies agencies' use of information from the Continuous National Survey. His respondents were able to document instrumental use of the results. Only rarely however, was research used immediately for such purposes as regulation writing. In a study by Patton et al. (1977) decision makers, evaluators, and government project monitors noted that evaluation results were frequently used in decision-making, but it was difficult for them to specify the link between the information and the decision. Weiss and Bucuvalas (1977) provide support for the utility of research for decision-making, if not instrumental use per se. In their study, officials in mental health reported that they found some evaluation results useful for choosing among alternative proposals for action.

Weiss (1977b) argues cogently for the importance of conceptual use in public policy and programs. Rich (1977) found that research was frequently used as background information on an issue. Caplan et al. (1975) concluded from their interviews with 204 administrators in federal government that information was often incorporated into officials' overall orientation toward a policy question. In Weiss and Bucuvalas's study, respondents said that research results had utility to change ways of thinking about an issue. Moreover, they rated as useful research findings that questioned existing policy. Such information may not be applied immediately, but may serve as part of the basis upon which policies are revised.

Conceptual use can lead to instrumental use at a later time. Rich's respondents frequently said that results would be used for specific policy decisions in the future. Whether future use occurred could not be established in his study. In Knorr's (1977) study of utilization by officials in the Austrian government, research was most often used as background information (43% of respondents). In many cases however, it served both as background and as input to specific decisions (19%).

Knorr (1977) provides the only quantified evidence for persuasive use, or as she termed it, "decision legitimative" use. Of her respondents, 11% said that they had used research to legitimate policies they had

intended to carry out in any case. Some of Patton et al.'s interviews (1977: 147) indicate persuasive use, however:

> Well, I think that all we did was probably speed up the process. . . . They knew that their performance was being criticized by various parts of the government and the private sector. . . . We were able to show just how poor a job they were doing.

DIFFICULTIES WITH CURRENT DEFINITIONS OF CATEGORIES

There are some problems with the above three definitions and ways they are made operational. For example, persuasive use has been discussed, to date, as though it were merely lip service to research (Pelz, 1978; Young and Comptois, 1979). Weiss (1977a) has pointed out, however, that use of results as "political ammunition" can be constructive and legitimate. For example, Bauman (1976) describes how the research on Health Maintenance Organizations (HMOs) was first used to persuade the Nixon Administration to adopt HMOs as part of its health policy, then used as ammunition when the debate over HMOs intensified. Fairweather et al. (1974), in their seminal work on adoption of innovations, describe an explicit phase of persuading mental hospitals to adopt an innovative technique. In other words, people do seriously discuss evaluations in persuasive utilization.

Some previous discussions of persuasive use have been unsatisfactory because they do not meet our "bottom-line" criteria. Using research to delay action, to allow policy makers to appear concerned about a problem, or to jockey for political position (Pelz, 1978; Weiss, 1977a) are not instances of utilization at all, if there is no evidence that evaluations were seriously discussed. Moreover, in these instances, the research process is used, not research information, for political purposes. The definition of utilization must remain consistent.

Current definitions of instrumental and conceptual utilization create problems as well. Given the wording of the definitions, people frequently confuse the ways in which these categories are made operational with the theoretical constructs. For example, the theoretical construct, "instrumental use," is application of information to a decision or problem. It is made operational by requiring that such application be documented. However, not all application of information to decisions is documented. Evidence of utilization within closed meetings, or a single individual's choice among alternatives, may exist nowhere in writing. Similarly, conceptual use may be amply documented—a background paper, for example, could identify changes in

agency thinking on an issue over time without setting any specific problem to be solved. It is important to clarify the difference between the construct and the way it is made operational, because some have erroneously equated instrumental use with the simplistic "impact" definition of utilization.

A second problem associated with the distinction between instrumental and conceptual utilization was addressed by Rich himself. He discovered that it was difficult for respondents to trace a specific decision back to particular sources of information (Pelz, 1978). Patton and his colleagues (1977) arrived at the same conclusion. Rein and White (1975) concur with these researchers that problems in government are defined gradually over time, and decisions are eventually reached on the basis of an integrated set of information from many sources. Under these circumstances it is difficult to determine where conceptual use ends and instrumental use begins.

A RECONCEPTUALIZATION OF UTILIZATION CATEGORIES

Pelz (1978) correctly notes that the categorization of specific instances of use is somewhat blurred, and that uses of research must often be categorized as "primarily instrumental" or "primarily conceptual," for example. Given current definitions, however, readers frequently assume that evaluation is used solely for specific decisions, or solely in people's thinking about a program. We can alleviate this confusion by viewing an evaluation as a "reference work," which as Boruch and Wortman (1979) and Weiss (1978) suggest, has the potential for being used and reused for many purposes. Categorizing uses then involves identifying the purposes that an evaluation serves at a particular time.

We suggest that cycles of bureaucratic decision-making and policy revision determine the type of use to which evaluations can be put. For example, instrumental use of information is not possible if there is no relevant decision pending. Congressional cycles are the clearest example. Florio et al. (1979) and Mitchell (1980) describe the uses of evaluations at different phases of these cycles. Other cycles include regulation writing, congressional oversight, and performance monitoring by agencies. One difficulty in categorizing uses according to policy cycles is that often these cycles do not follow clearly defined stages (Rein and White, 1975). However, if the methodology for determining use is longitudinal, rather than retrospective, it should be possible to identify

the policy activities that are emphasized at any particular time and determine the influence, if any, of evaluation results on these activities.

METHODOLOGICAL ISSUES

Problems of method abound in the study of utilization. These arise because of types of research strategies made necessary by the subject matter, because of the low priority given to documentation of utilization, and because of misconceptions of utilization itself. Most research on utilization has relied either on the case study method or on policy makers' statements in interviews and surveys. Case studies include Alkin and associates' (1979) descriptions of the utilization of evaluations in five school districts, and Menges's (1978) history of the use of evaluations in policy on equal educational opportunity. These studies are strongest in their ability to plausibly link evaluation findings and evidence of use. Their weakness lies in the lack of generality and confounding of situational factors. However, one method with great potential is content analysis of a sample of cases that are representative of evaluations in a policy sector.

Interviews and surveys have varied in content and in the underlying definition of utilization. For example, some have focused on particular groups of officials and their use of applied research, in general (Caplan et al., 1975; Florio et al., 1979; Knorr, 1977; Weinberg, 1979). Others studied the use of particular research projects (Patton et al., 1977; Rich, 1977). Weiss and Bucuvalas (1977) and Nielsen (1975) studied the utility of applied social research, not utilization itself. Brown et al. (1978) and Eaton (1969) examined reactions to evaluations. Of all these studies, only those by Caplan et al., Rich, Weiss and Bucuvalas, and Knorr explicitly differentiated any of the major types of utilization given in our definition. Differences among the respondents interviewed in these studies probably influenced the findings. As shown by Resnick et al. (1979), people with different responsibilities use evaluations differently. Some studies have interviewed a heterogeneous group of users (David, 1978; Lyon et al., 1978; Resnick et al., 1979; Weiss and Bucuvalas, 1977). However, others have interviewed only administrators (Caplan et al., 1975; Knorr, 1977; Nielsen, 1975) or practitioners (Brown et al., 1978; Eaton, 1969), or congressional staff (Florio et al., 1979; Weinberg, 1979).

Case studies, interviews, and questionnaires all suffer from at least four major problems in the study of utilization. First, it is difficult to document that utilization occurs, because evaluations are frequently used informally (Pelz, 1978; Rein and White, 1975; Rich, 1979) and because a substantial time may lapse between utilization and the study of utilization. Given officials' faulty memories, retrospective research may be biased in favor of a few dramatic instances of use, rather than frequent but modest ones. A related problem involves demonstrating that change, at any level, was caused at least in part by evaluations. Rich (1977) is one of the very few researchers who has been able to document utilization.

A third problem is the question of base rates for comparison. Until recently, we believed that the base rate for utilization was very low. We are learning that the fault may lie with our measures, not with evaluation. A fourth problem involves the unit of analysis: What is an instance of utilization? It is necessary to quantify utilization if we are to show that it can be enhanced. Most interview respondents treat impact as the unit of analysis, but as Pelz (1978) has shown, several types of utilization may be involved in producing impact. An example shows how important the unit of analysis can be. Content analysis of House and Senate Committee reports reveals that an evaluation sponsored by NIE was used for 22 separate amendments to compensatory education legislation in 1978 (Leviton and Boruch, 1980). It would be misleading to claim a single "use" of the NIE evaluation in these hearings.

Because the existing studies share these flaws, our review of variables affecting utilization can only be tentative. It is for this reason that we supplement studies of utilization with other information that may not share these biases in the review of variables affecting utilization.

VARIABLES AFFECTING UTILIZATION

Five major clusters of variables are consistently related to utilization between evaluators and users, information processing by users, credibility of evaluations, and user involvement. Tables 1 through 5 summarize the variables to be described, for each of the five clusters. A plus sign (+) appears to the left of a variable if, as it is described in the table, it enhances utilization. A minus sign (–) appears if the variable detracts from utilization. A plus and minus (±) indicate that the variable

TABLE 1
Relevance

+A.   Evaluations address clients' needs

B.   Policy makers' needs
    + Resource allocation information
    + Implementation information
    + Information about overall effectiveness

C.   Program managers' needs
    + Implementation information
    + Information about effectiveness of program elements
    − Information about overall effectiveness takes lower priority

D.   Timeliness
    ± Probable interaction with type of use
    ± Interaction with stage of policy cycle
    + Lead time for consideration of evidence

interacts with other variables, enhancing utilization in some cases and detracting or having no effect in others. Such interactions help explain contradictory findings, such as those involving quality and timeliness. Those variables posing a danger of misutilization are labelled as such.

RELEVANCE

Many writers have noted that research would be used more frequently if it had higher relevance to policy or program concerns (Averch, 1975; Banta and Bauman, 1976; Cox, 1977; Williams and Wysong, 1975). Lynn (1978) notes, however, that relevance is difficult to achieve. The difficulties lie in two areas: addressing clients' needs and timeliness of information (see Table 1).

Several studies relate clients' needs to utilization. Nielsen (1975) compared program mangers' needs for information with the information they actually obtained through evaluations. David (1978) found low utilization by school district audiences in a survey of local evaluations of compensatory education. Both researchers concluded that low relevance of evaluation content caused low utilization. Since a very low rate of utilization was obtained, however, these studies could not demonstrate that high relevance would increase utilization. Resnick et al. (1979) were able to show this. Evaluations that were relevant to the

needs of a particular audience were used more frequently than less relevant evaluations. In this study, evaluations dealing with resource allocations were used by school boards and by superintendents; evaluations involving curriculum changes were used by superintendents, teachers and program directors.

The study by Resnick et al. demonstrates the divergent uses by policy makers and by program staff. At the state and federal levels, a similar pattern emerges (Pelz, 1979; White and Murnaghan, 1973). Both program and policy audiences find practical implementation information useful (Nielsen, 1975; Weiss and Bucuvalas, 1977). While policy makers in Weiss and Bucuvalas's study rated as useful overall statements of the effectiveness of programs, program managers in the Nielsen study did not. Program managers preferred information about the effectiveness of different program elements. Other evidence indicates that studies of overall effectiveness do not have priority with most program managers and with those who deliver services (Brickell, 1974; Wholey et al., 1970).

A second aspect of relevance is timeliness. Many writers have heavily emphasized the importance of submitting results in time for a decision (Banta and Bauman, 1976; Falcone and Jaeger, 1976; Guba, 1975). In contrast, two studies have concluded that timeliness is not consistently important in getting research used. Weiss and Bucuvalas (1977) asked their respondents to rate the extent to which the attribute "on time for a pending decision" applied to the studies described to them. Timeliness accounted for less than 1% of the variance in predicting perceived usefulness, when other variables were taken into account. This finding may not reflect the importance of timeliness in practice, because officials in this study may not have felt any urgency over hypothetical applications of research information. However, Patton et al. (1977) interviewed the most directly relevant dicision makers for particular evaluations. In retrospect, these respondents reported that getting reports in on time for a specific decision had not been important in getting the results used.

How can the insistent emphasis on timeliness be reconciled to these findings? Young and Comptois (1979) suggest that timeliness may be important for short-term instrumental use but less so for conceptual use (which is heavily emphasized in both studies above). Florio et al. (1979) demonstrate the importance of delivering evaluations to Congress during particular phases of the legislative cycle—evaluations cannot be used if the relevant hearings are over. In fact, lead time appears to be necessary if evaluations are to influence thinking and actions of legislators (Hill, 1980; Mitchell, 1980). Timeliness may affect use in

TABLE 2
Communication

+A.	Direct communication of users' needs for evaluation
−B.	Communication in bureaucratic hierarchies
C.	Dissemination

  + Direct dissemination
  − Information left out of summaries
  − Distortions or misstatements in summaries − misutilization possible
  − Negative information suppressed − misutilization possible
  + Dissemination networks
  − Dissemination between agencies
  + "Invisible colleges" of specialists

other cyclic activities such as regulation writing (Millsap, 1978). Because several writers have noted a trade-off between quality and timeliness of evaluations (Rein and White, 1975), an accurate assessment is desirable of circumstances under which timeliness is important.

COMMUNICATION

Two studies demonstrate the importance of good communications between producers of evaluation and potential users. Windle and Bates (1974) examined 15 evaluations conducted for NIMH. The contractor and NIMH staff communicated more frequently in the more useful studies. Glaser and Taylor (1973) compared 5 "successful" and 5 "unsuccessful" NIMH research projects. Successful studies were used more frequently in documentable ways. Glaser and Taylor determined that close communication among consumers and researchers distinguished successful studies at each stage of development of a project. An outstanding example of the effects of good communication is presented by Hill (1980). Variables affecting communication are shown in Table 2.

Unfortunately, communications within bureaucracies tend to be obstructed. Downs (1967) concluded, after extensive study, that "middlemen" in the bureaucratic hierarchy selectively pass information between upper and lower echelons. Information may thus be eliminated or distorted. Communication through the bureaucratic hierarchy can adversely affect utility, because the evaluator may fail to gain a complete, unbiased idea of users' needs. Davis and Salasin (1978) give an example in which one "middleman" changed the entire focus of an

evaluation with the change of one phrase in the directive. The hierarchy affects dissemination, because valuable information may never get to potential users. For example, administrators in the higher echelons usually read executive summaries of reports or have subordinates summarize reports for them (Brickell, 1974). Information and qualifications may be left out (Cook and Pollard, 1977). The quality of summaries may also suffer due to deliberate distortions or to inadvertent misstatements (Datta, 1977). Both research (Eaton, 1969) and informal observation (Guba, 1975) reveal a tendency to leave negative information out of reports.

Weiss (1978) notes that formal dissemination networks are somewhat rare in bureaucracies, a point coroborated by the MITRE conference on utilization (Chelimsky, 1977). Caplan et al. (1975) noted a strong tendency to use only information generated by administrators' own agency. This finding is explained in part by the lack of dissemination between agencies (Weiss, 1978). However, information is passed informally between agencies if users form part of an "invisible college" of specialists. Rich (1979) studied a group of economists in various agencies who worked in the area of unemployment insurance. Their written policy positions were modified by, but did not cite, information they received informally from outside the agency. Eaton (1969) also found a preference for informal communication of research findings outside the practitioners' own programs.

INFORMATION PROCESSING

A criterion for utilization is that relevance be recognized and the findings translated into their implications. Variables affecting this task are shown in Table 3. Evaluations differ from some other social research in that users are usually aware of the relevance of evaluations to policies and programs. Even so, it frequently takes work to determine the relevance of many evaluations. For example, goals may not be clearly specified (Wholey et al., 1970). It is difficult to determine the federal significance of local evaluations, received by some agencies (Davis et al., 1977). Finally, in even the most specific consumer-oriented evaluations, information may emerge that users did not anticipate, as when negative side effects of programs are discovered (Cook and Pollard, 1977).

The way evaluation is presented to users affects their comprehension and thus the extent of use. Several writers suggest that readable reports are utilized more (Agarwala-Rogers, 1977; Windle and Bates, 1974). Some writers suggest that explicit recommendations in evaluation

TABLE 3
Information Processing

A. Awareness of relevance
    + Evaluations set up to answer specific questions – probable effect
    – Goals not clearly specified – probable effect
    – Federal relevance of local evaluations – probable effect
    ± Unanticipated information – probable interaction with user's abilities

B. Presentation of information
    + Clear presentation
    – Jargon increases difficulty

C. Information processing style of administrators
    – Attention to relevant information only
    – Little scrutiny of quality – misutilization possible
    + Verbal communication
    – Lack of familiarity with methods – misutilization possible
    ± Preference for qualitative information – interaction with
      evaluation content
    – Difference between administrator and academic styles

reports may enhance utilization (Datta, 1977; Davis et al., 1977). In general, clear communication of knowledge has been found to affect diffusion (Zaltman et al., 1973). Brown et al. (1978) found that although technical jargon did not influence ratings of the usefulness of evaluations, ratings of difficulty were affected. Utilization may be impaired if the incentive to read difficult reports is low.

Cox (1977) has adapted Mintzberg's (1973) research on the activities of managers to characterize the information processing style of administrators of social programs. The implications of this style are that (1) only findings relevant to managers' needs will receive much attention; (2) critical scrutiny of quality is unlikely; and (3) close verbal communication will enhance utilization, because managers prefer to be informed on a continuing basis. Cox mentions that, for many managers, evaluation is an unfamiliar tool. It may be possible to generalize this statement for many types of users. Weinberg's (1979) interview study describes congressional staffers' ignorance of common flaws in social research. The respondents of Caplan et al. (1975) rated observations of real life situations as more reliable than laboratory studies. One consequence of this lack of expertise may be misutilization.

Administrators prefer qualitative information to the quantitative data that evaluators frequently supply. This argues for combining qualitative and quantitative methods. Officials in Caplan's (1975: 18) survey most frequently used "soft knowledge" (nonresearch based, qualitative and couched in lay language) as opposed to "hard knowledge" (research-based, usually quantitative, and couched in scientific language). Nielsen's (1975) study of program managers corroborates users' preference for qualitative information. Finally, Weinberg (1979) and Florio et al. (1979) describe congressional staff's extensive use of vivid examples rather than quantitative information.

The difference in styles between researchers and administrators may affect utilization. Caplan and his colleagues found little contact between administrators and the academic community. The values of the two groups have been found to differ (Tiffany et al., 1971). Some have suggested that "knowledge brokers" can bridge the gap between these two ways of looking at policy and programs (Sundquist, 1978). Others have suggested that increased trust may grow out of increased familiarity (Caplan, 1979).

CREDIBILITY

Just as an evaluator uses multiple indicators of an outcome, an administrator has multiple indicators of the faith that can be placed in information (Downs, 1967). These indicators are summarized in Table 4. Administrators have other information available to them on issues besides a single evaluation (Caplan et al., 1975). This information includes, for example, the day-to-day experience of program functioning (Weiss, 1979; Weiss and Bucuvalas, 1977). Surveys of policy makers reveal that if evaluation results are surprising in light of these other sources of information, or in light of client's expectations or intuition, they are less likely to be used (Caplan et al., 1975; Patton et al., 1977; Weiss and Bucuvalas, 1977). Evaluations are used to corroborate other information, as shown by the finding of Patton et al. (1977) that evaluations are used to increase certainty about existing knowledge and decisions. Rich (1977) found that research reports were used in aggregates, again because the administrator gains a convergence of knowledge.

Certain biases of administrators may make them more or less receptive to research. Caplan et al. (1975) found that heavy users had a higher opinion of the objectivity of research and more frequently

TABLE 4
Credibility

A. Comparison with other information
    − Contradiction between research and other information
    − Contradication of users' intuition or expectations
    + Corroborating information
    + Use of research in aggregates

B. Preconceptions of users
    + High opinion of research
    + Reliance on research over intuition
    + User has medical degree
    − User has law degree − (persuasive use likely)
    + User is in government temporarily

C. Credibility of evaluation producer
    − Probable effect of suspected cooptation
    − Probable effect on use outside agency

D. Quality
    − Low perceived quality
    − Low quality − misinformation possible
    ± Interaction with type of use − convincing others
    + Methodology not easily assailed

endorsed the statement that intuition was not as good a source of information as research. Educational background made a difference: MDs used research most often, people with law degrees, least often. A law degree was, moreover, associated with using research in the service of political expediency, a form of persuasive use, as opposed to other types. Heavy users of research were likely to be in government only temporarily, reinforcing Downs's (1967) finding that bureaucrats with short tenure are less likely to resist change.

Credibility of the producer of an evaluation is likely to be important. Expertise of the evaluator did not affect practitioners' trust of the findings in a study by Brown et al. (1978), although it may influence other audiences. More important may be suspicion that the researchers have been coopted or have suppressed information. Guba (1975) warns that too many evaluators are willing to produce the data their clients want. Brickell (1978) describes the monetary pressures exerted by clients of evaluations. Windle and Bates (1974) give examples of the low eredibility that evaluations can have with policy makers when coopta-

tion by the sponsor is assumed. From this anecdotal evidence, we suggest that the trustworthiness of an evaluation is likely to be a more important issue when the evaluation is disseminated outside the program or agency that sponsors it.

Although a number of writers have advocated higher methodological quality as a means of increasing utilization (Wholey et al., 1970), quality has not been found to consistently affect utilization (Patton et al., 1977). If quality does influence use, it is likely to do so primarily through increased trust that the findings are an accurate picture of the program. Administrators interviewed by Caplan et al. (1975) frequently singled out shoddy program evaluation as discrediting the objectivity of evaluations in their eyes. Poor quality is related to misutilization of findings (Cook and Polard, 1977).

Methodological quality is most likely to influence utilization when it is essential that the data be convincing. Weiss and Bucuvalas (1977) asked officials to rate both the quality of research information and the likelihood that it would be used in a variety of ways. Research quality was related to only two uses: mobilizing support for a position and changing ways of thinking about an issue. Weiss and Bucuvalas point out that in both cases, resistance to the information must be overcome. Since evaluations are frequently attacked on the basis of their methodology (Brickell, 1978; Davis and Salasin, 1975), quality adds to the persuasiveness of a position.

USER INVOLVEMENT AND ADVOCACY

In a political context, advocacy is expected (Brandl, 1978). However, advocates of evaluation information must compete with advocates of other kinds. The relationship between evaluations and advocacy is a large part of the political context in which evaluations are conducted. Both Patton et al. (1977) and Caplan and his colleagues (1975) found extensive agreement among their respondents that this context is essential to understand utilization. Aspects of this relationship are shown in Table 5.

Fairweather et al. (1974) and Glaser (1976) present evidence that persistent advocacy by a key individual is essential in getting research findings used. Patton and his colleagues (1977) concluded that the commitment of an individual decision maker determines whether evaluations are used. Evaluators and policy makers in a conference held by the MITRE Corporation (Chelimsky, 1977) corroborated this point.

TABLE 5
User Involvement and Advocacy

A.	Commitment to Evaluations
	+ High interest and commitment by user
	+ Advocacy of evaluation information
	± Probable interaction of user commitment with policy change
	± Probable interaction of user commitment with turnover in users
B.	Advocacy of Programs and Policies
	± Interaction with direction of findings
	+ Less visible evaluations – probable effect on impact
	+ Smaller constituencies – probable on impact
	± Probable interaction of outcome evaluation with time

A study of the demand for evaluations also indicates that specific individuals are the primary consumers (Schmidt et al., 1977).

The impact of specific individuals on organizations is seldom as powerful as in the case of research utilization. For example, Pfeffer and Salancik (1978) found that the actions of an organizational leader explained only 10% of the variance in behavior of organizations they studied. Why the unusually strong effects of committed individuals in the use of evaluations? One set of explanations for the importance of a committed individual deals with the types of utilization involved. Patton et al. found only one case of major impact. The majority of uses in this study may have been the sole responsibility of the committed individual. Alternatively, the committed individual's power may have dictated utilization by others. Persuasive use by this individual may also have convinced others.

A second set of explanations have to do with communication in bureaucracies. By taking responsibility for getting information to users, the committed individual may bypass middlemen who distort or eliminate findings. The individual may also clarify the goals for evaluation by effectively communicating personal needs for information. Finally, the individual may have an effect through the ability to translate the findings into their policy implications (see Hill, 1980).

Patton's (1978) conclusion from this work is that utilization is enhanced by gearing evaluations to the particular individuals who are receptive to them. Note however, that the advantages of this approach are limited by two bureaucratic realities: rapid policy change and the rapid turnover among upper level administrators in government. Rapid policy changes limit this approach, because administrators often

respond to short-term pressures for information (Chelimsky, 1977; Mechanic, 1978; Weiss, 1978). Evaluators could conscientiously address these short-term needs, only to find that the focus of policy had shifted to a new set of issues. The length of time it takes to do an evaluation may also be such that the potential user has taken another job. Evaluations may be suited to the needs of new potential users, but not if the subject matter is geared to idiosyncratic needs of one or two people.

How then, can one take advantage of Patton's insight? A marketing strategy based on the role of the user in government might be an effective compromise. It should be possible to identify the information needs associated with a position in government. For example, the needs of the Medicare Bureau Director remain somewhat the same, regardless of the person holding the job at a particular time. Constituency analyses, such as the one performed by Brickell (1974), provide information about such needs. This approach could be combined with a "feeling of ownership" to enhance use.

People more frequently will be advocates of programs and policies than advocates of an evaluation. Potential users' investment—emotional or material—in a program or policy affects their acceptance of findings. While the conclusions of a study, positive or negative, have not been consistently found to affect utilization (Patton et al., 1977), the direction of findings interacts with the position taken by potential users. Advocates of a program may become advocates of evaluations that support their position. On the other hand, evaluations that run counter to advocacy will be attacked. For example, Carter (1971) described case studies in which administrators did not accept evaluation findings that went against their own interests. Davis and Salasin (1975) reported a study of newspaper coverage of evaluations. Almost invariably, the program manager attacked the validity of negative findings. Campbell (1969) notes that a material interest is not necessary for administrators to become wedded to an advocacy position. Administrators may become overly committed to past decisions, such that they ignore or attack evaluation findings that indicate that these decisions were wrong. Staw has conducted a series of experiments that support Campbell's observation (Fox and Staw, 1979; Staw, 1976; Staw and Fox, 1977; Staw and Ross, 1978). A public commitment is the crucial variable, not an attitude as such (Kiesler, 1971).

Because large programs have large constituencies, evaluations of their outcome may not be used very much. However, marginal changes occur within such programs, over which evaluation may have a degree of influence (George Myeske, personal communication). Certain kinds

of evaluations may be more likely to be used than others: Brandl (1978) suggests that evaluations of implementation or of demonstration programs may be used more, because they have less visibility and challenge fewer constituents than do outcome evaluations. Rein and White (1975) note that the crucial variable influencing use of outcome evaluations may be the passage of time. However, it is likely that all these researchers are really discussing impact—conceptual use of evaluations may be immediate and extensive.

SUMMARY

This review has attempted to clarify existing conceptions of utilization. It has suggested improvements in the methods of detecting use. Finally, it described five major clusters of variables that affect utilization and suggested hypotheses about the reasons for their effects. By studying utilization, we can improve our methods, because utilization is intimately associated with the plan for evaluation (Windle and Bates, 1974). We can also exercise fate control as professionals, by anticipating the likely uses of our work and guiding these where it is appropriate and responsible to do so.

REFERENCES

AGARWALA-ROGERS, R. (1977) "Why is evaluation research not utilized?" in M. Guttentag (ed.) Evaluation Studies Review Annual, Vol. 2. Beverly Hills: Sage.

ALKIN, M. C., R. DAILLAK, and P. WHITE (1979) Using Evaluations: Does Evaluation Make a Difference? Beverly Hills: Sage.

AVERCH, H. (1975) "Notes on improving research utility." Inquiry, 12, 3 (September): 231-234.

BANTA, H. D. and P. BAUMAN (1976) "Health services research and health policy." J. of Community Health 2, 2 (Winter): 121-132.

BAUMAN, P. (1976) "The formulation and evolution of the health maintenance organization policy, 1970-1973." Social Sci. and Medicine 10, 3/4 (March/April): 129-142.

BORUCH, R. F. and P. M. WORTMAN (1979) "Implications of educational evaluation for evaluation policy," in D. Berliner (ed.) Review of Research in Education (Vol. 7). Washington, DC: American Educational Research Association.

BRANDL, J. E. (1978) "Evaluation and politics." Evaluation (Special Issue): 6-7.

BRICKELL, H. M. (1978) "The influence of external political factors on the role and methodology of evaluation," in T. D. Cook et al. (eds.) Evaluation Studies Review Annual (Vol. 3). Beverly Hills: Sage.

——— (1974) Data for Decisions: An Analysis of Evaluation Data Needed by Decision-Makers in Educational Programs. New York: Policy Studies in Education.

BROWN, R. D., L. A. BRASKAMP, and D. L. NEWMAN (1978) "Evaluator credibility as a function of report style." Evaluation Q. 2, 2 (May): 331-341.

CAMPBELL, D. T. (1969) "Reforms as Experiments." Amer. Psychologist 24, 4 (April): 409-429.

CAPLAN, N. (1979) "The two-communities theory and knowledge utilization." Amer. Behavioral Scientist 22, 3 (January/February): 459-469.

——— A. MORRISON, and R. STAMBAUGH (1975) The Use of Social Science Knowledge in Policy Decisions at the National Level. Ann Arbor, MI: Institute for Social Research.

CARTER, R. K. (1971) "Client reactions to negative findings and the latent conservative function of evaluation studies." Amer. Sociologist 6, 2 (May): 118-124.

CHELIMSKY, E. (1977) A Symposium on the Use of Evaluation by Federal Agencies, (Vol. II). Washington, DC: National Institute of Law Enforcement and Criminal Justice and Metrek, a division of the MITRE Corporation.

COOK, T. D. (1978) "Introduction," in T. D. Cook et al. (eds.) Evaluation Studies Review Annual. Beverly Hills: Sage.

——— and W. E. POLLARD (1977) "Guidelines: how to recognize and avoid some common problems of mis-utilization of evaluation research findings." Evaluation 4: 161-164.

COOK, T. D., J. LEVINSON-ROSE, and W. B. POLLARD (1980) "The misutilization of evaluation research: some pitfalls of definition." Knowledge 1, 4 (June): 477-498.

COX, G. B. (1977) "Managerial style: implications for the utilization of program evaluation information." Evaluation Q. 1, 4 (December): 499-508.

DATTA, L. E. (1977) "Does it work when it has been tried? and half full or half empty?" in M. Guttentag (ed.) Evaluation Studies Review Annual (Vol. 2). Beverly Hills: Sage.

DAVID, J. L. (1978) Local Uses of Title I Evaluations. Report prepared for the Office of the Assistant Secretary for Planning and Evaluation, Department of Health, Education and Welfare. Menlo Park, CA: SRI International.

DAVIS, H. R. and S. E. SALASIN (1978) "Strengthening the contribution of social R&D to policy making," in L. E. Lynn (ed.) Knowledge and Policy: The Uncertain Connection. Washington DC: National Research Council.

——— (1975) "The utilization of evaluation," in E. K. Struening and M. Guttentag (eds.) Handbook of Evaluation Research (Vol. 1). Beverly Hills: Sage.

DAVIS, H. R., C. WINDLE, and S. S. SCHARFSTEIN (1977) "Developing guidelines for program evaluation in community mental health centers." Evaluation 4 (Special Issue): 25-29.

DOBSON, A., J. G. GREER, R. H. CARLSON, F. A. DAVIS, L. E. KUCKEN, B. J. STEINHARDT, T. P. FERRY, and G. S. ADLER (1978) "PSRO's: their current status and their impact to date." Inquiry 15, 2 (June): 113-128.

DOWNS, A. (1967) Inside Bureaucracy. Boston: Little, Brown.

EATON, J. W. (1969) "Symbolic and substantive evaluation research," in H. C. Schulberg et al. (eds.) Program Evaluation in the Health Fields. New York: Behavioral Publications.

FAIRWEATHER, G. W., D. H. SANDERS, and L. G. TORNATSKY (1974) Creating Change in Mental Health Organizations. New York: Pergamon Press.

FALCONE, D. and B. J. JAEGER, (1976) "The policy effectiveness of health services research: a reconsideration." J. of Community Health 2, 1 (Fall): 36-51.

FLORIO, D. H., M. M. BEHRMAN, and D. L. GOLTZ (1979) "What do policy makers think of educational research and evaluation? Or do they?" Educ. Evaluation & Policy Analysis 1: 61-88.

FOX, F. W. and B. M. STAW (1979) "The trapped administrator: effects of job security and policy resistance upon commitment to a course of action." Evanston, IL: Northwestern University. (unpublished)

GLASER, E. M. (1976) Putting Knowledge to Use: A Distillation of the Literature Regarding Knowledge Transfer and Change. Beverly Hills: Human Interaction Research Institute.

——— and S. H. TAYLOR (1973) "Factors influencing the success of applied research." Amer. Psychologist 28, 2 (February): 140-146.

GUBA, E. G. (1975) "Problems in utilizing the results of evaluation." J. of Research and Development in Education 8, 3 (Spring): 42-54.

HILL, P. (1980) "Evaluating education programs for federal policy makers: lessons from the NIE Compensatory Education Study," in J. Pincus (ed.) Educational Evaluation in the Public Policy Setting. Santa Monica, CA: Rand Corporation.

Huron Institute (1979) Proposal to National Institute of Education. Cambridge, MA: Author.

KIESLER, C. A. (1971) The Psychology of Commitment. New York: Academic.

KNORR, K. D. (1977) "Policymakers' use of social science knowledge: symbolic or instrumental?" in C. H. Weiss (ed.) Using Social Science Research in Public Policy Making. Lexington, MA: Lexington Books.

LARSEN, J. K. and P. D. WERNER (1981) "Measuring utilization of mental health program consultation," in J. A. Ciarlo (ed.) Utilizing Evaluation. Beverly Hills: Sage.

LEVITON, L. C. and R. F. BORUCH (1980) "Illustrative case studies," in R. F. Boruch and D. S. Cordray (eds.) An Appraisal of Educational Program Evaluations: Federal, State, and Local Levels. Washington, DC: U.S. Department of Education.

LYNN, L. E. Jr. (1978) "The question of relevance," in L. E. Lynn (ed.) Knowledge and Policy: The Uncertain Connection. Washington DC: National Research Council.

LYON, C. D., L. DOSCHER, P. MCGRANAHAN, and R. WILLIAMS (1978) Evaluation and School Districts. Los Angeles: Center for the Study of Evaluation.

MECHANIC, D. (1978) "Prospects and problems in health services research." Milbank Memorial Fund Q. 56, 2 (Spring): 127-139.

MENGES, C. C. (1978) Knowledge and Action: The Rise of Social Science Evaluations in Decisions on Equal Educational Opportunity, 1970-1973. Washington, DC: National Institute of Education.

MILLSAP, M. A. (1978) "The uses of evaluation in federal regulation writing: a case study." Presented at Evaluation Research Society Convention, Washington, D.C.

MINTZBERG, H. (1973) The Nature of Managerial Work. New York: Harper & Row.

MITCHELL, D. E. (1980) "Social science impact on legislative decision making: process and subs.ance." Presented at the annual meeting of the American Educational Research Association.

MUSHKIN, S. J. (1973) "Evaluations: use with caution." Evaluation 1, 2: 31-35.

NIELSEN, V. G. (1975) "Why evaluation does not improve program effectiveness." Policy Studies J. 3, 4 (Summer): 385-389.

PATTON, M. Q. (1978) Utilization-Focused Evaluation. Beverly Hills: Sage.

——— P. S. GRIMES, K. M. GUTHRIE, N. J. BRENNAN, B. D. FRENCH and D. A. BLYTH (1977) "In search of impact: an analysis of the utilization of federal health

evaluation research," in C. H. Weiss (ed.) Using Social Research in Public Policy Making. Lexington, MA: Lexington Books.

PELZ, D. C. (1979) Knowledge Utilization in Criminal Justice. A Research Proposal Submitted to the National Institute for Law Enforcement and Criminal Justice, LEAA. Ann Arbor, MI: Institute for Social Research.

——— (1978) "Some expanded perspectives on use of social science in public policy," in J. M. Yinger and S. J. Cutler (eds.) Major Social Issues: A Multidisciplinary View. New York: Macmillan.

PFEFFER, J. and G. R. SALANCIK (1978) The External Control of Organizations: A Resource Dependent Perspective. New York: Harper & Row.

REIN, M. and S. H. WHITE (1975) "Can policy research help policy?" The Public Interest 49 (Fall): 119-136.

RESNICK, S. M., C. O'REILLY, and A. MAJCHRZAK "Utilization: different users different uses." Presented at Evaluation Research Society Convention, Minneapolis.

RICH, R. F. (1979) "Problem-solving and evaluation research: unemployment insurance policy," in R. F. Rich (ed.) Translating Evaluation into Policy. Beverly Hills: Sage.

——— (1977) "Uses of social science information by federal bureaucrats: knowledge for action versus knowledge for understanding," in C. H. Weiss (ed.) Using Social Research in Public Policy Making. Lexington, MA: Lexington Books.

SCHMIDT, R. E., P. HORST, J. W. SCANLON, and J. S. WHOLEY (1977) Serving the Federal Evaluation Market. Washington, DC: Urban Institute.

STAW, B. M. (1976) "Knee-deep in the big muddy: a study of escalating commitment to a chosen course of action." Organizational Behavior and Human Performance 16, 1: 27-44.

——— and F. V. FOX (1977) "Escalation: the determinants of commitment to a chosen course of action." Human Relations 30, 5 (May): 431-450.

STAW, B. M. and J. ROSS (1978) "Commitment to a policy decision: a multitheoretical perspective." Admin. Sci. Q. 23, 1 (March): 40-64.

SUNDQUIST, J. L. (1978) "Research brokerage: the weak link," in L. E. Lynn (ed.) Knowledge and Policy: The Uncertain Connection. Washington DC: National Research Council.

TALLMADGE, G. K. (1977) Ideabook: The Joint Dissemination Review Panel. Washington DC: U.S. Office of Education and National Institute of Education.

TIFFANY, D. W., M. PHYLLIS and J. R. COWAN (1971) "A source of problems between social science knowledge and practice." J. of Human Relations 19, 2: 239-250.

WEINBERG, H. (1979) "The role of cognitive biases in the interpretation of program evaluations." Dissertation in progress, Yale University.

WEISS, C. H. (1978) "Improving the linkage between social research and public policy," in L. E. Lynn (ed.) Knowledge and Policy: The Uncertain Connection. Washington DC: National Research Council.

——— (1977a) "Introduction," in C. H. Weiss (ed.) Using Social Research in Public Policy Making. Lexington, MA: Lexington Books.

——— (1977b) "Research for policy's sake: the enlightenment function of social research." Policy Analysis 3, 4 (Fall): 531-545.

——— (1973) "Where politics and evaluation meet." Evaluation 1, 3: 37-45.

——— and M. J. BUCUVALAS (1977) "The challenge of social research to decision-making," in C. H. Weiss (ed.) Using Social Science Research in Public Policy Making. Lexington, MA: Lexington Books.

WEISS, J. A. (1979) "Access to influence: some effects of policy sector on the use of social science." Amer. Behavioral Scientist 22, 3 (January/February): 437-458.

WHITE, K. L. and J. H. MURNAGHAN (1973) "Health care policy formation: analysis, information and research." Int. J. of Health Services 3, 1 (Winter): 81-91.

WHOLEY, J. S., J. W. SCANLON, H. G. DUFFY, J. S. FUKUMOTO, and L. M. VOGT (1969) Federal Evaluation Policy. Washington, DC: Urban Institute.

WILLIAMS, S. R. and J. A. WYSONG (1975) "The uses of research in national health policy: an assessment and agenda." Medical Care 13, 3 (March): 256-267.

WINDLE, C. and P. BATES (1974) "Evaluating program evaluation: a suggested approach," in P. O. Davidson et al. (eds.) Evaluation of Behavioral Programs. Champaign, IL: Research Press.

YOUNG, C. J. and J. COMPTOIS (1979) "Increasing congressional utilization of evaluation," in F. Zweig (ed.) Evaluation in Legislation. Beverly Hills: Sage.

ZALTMAN, G., R. DUNCAN, and J. HOLBECK (1973) Innovations and Organizations. New York: John Wiley.

Laura C. Leviton is Assistant Professor at the Health Policy Institute, University of Pittsburgh. She has studied the uses made of evaluations by HHS's Health Care Financing Administration and by the Department of Education. Her special interests are research and evaluation for health policy and programs.

Edward F.X. Hughes is Director of the Center for Health Services and Policy Research, Northwestern University. He has written in the areas of health policy, health economics, and surgical manpower.