Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions

GARY T. HENRY AND MELVIN M. MARK

ABSTRACT

Although use is a core construct in the field of evaluation, neither the change processes through which evaluation affects attitudes, beliefs, and actions, nor the interim outcomes that lie between the evaluation and its ultimate goal—social betterment—have been sufficiently developed. We draw a number of these change mechanisms, such as justification, persuasion, and policy diffusion, from the social science research literature, and organize them into a framework that has three levels: individual, interpersonal, and collective. We illustrate how these change processes can be linked together to form "pathways" or working hypotheses that link evaluation processes to outcomes that move us along the road toward the goal of social betterment. In addition, we join with Kirkhart (2000) in moving beyond use, to focus our thinking on evaluation influence. Influence, combined with the set of mechanisms and interim outcomes presented here, offers a better way for thinking about, communicating, and adding to the evidence base about the consequences of evaluation and the relationship of evaluation to social betterment.

Use is a core construct in the field of evaluation.[1] Most if not all evaluators strive to have their evaluations used. Many if not most evaluators accept the idea that, at least in part, the merit of their work—the success or failure of their evaluation efforts—can be judged in terms of

Gary Henry • Professor, Andrew Young School of Policy Studies, Georgia State University, P.O. Box 4039, Atlanta, GA 30302-4039, USA; Tel: (1) 404-651-2343; E-mail: gthenry@gsu.edu.

whether and how an evaluation is used. Indeed, Preskill and Caracelli (1997, p. 215) report that 99% of the evaluators surveyed agreed or strongly agreed that both providing information for decision-making *and* improving programs—classic aspects of evaluation use—were purposes for conducting evaluations, far more than agreed with any other purpose.

Use is generally understood to refer to a direct action that has occurred as a result of an evaluation (i.e., instrumental use) or to something that is newly learned about a program, its participants, its operations, or outcomes through an evaluation (i.e., conceptual use). The action or learning can take place as a result of evaluation findings, or as a result of participation in evaluation procedures (i.e., process use). In some cases, the concept of use is used to refer to waving the flag of evaluation to claim a rational basis for action (or inaction), or to justify pre-existing positions (i.e., symbolic use). Use can be premeditated, as exemplified by the mantra of utilization focused evaluation, "intended use by intended users" (Patton, 1997), or it can emerge as an evaluation unfolds, findings are generated, and opportunities arise.

Use has served the field of evaluation as both a motivating concept and a focus of theory, research, and debate. For example, use and its synonym, utilization, have been hotly debated by leaders in the field of evaluation, as in the debates of Patton and Weiss (Patton, 1988; Weiss, 1988a, 1988b). Use has also been a focus of research. The fundamental taxonomy of use, which includes instrumental, conceptual, and symbolic forms of use, draws heavily upon the research on evaluation use of the mid-1970s to very early-1980s (Alkin, Daillak, & White, 1979; Caplan, 1977; King, Thompson, & Pechman, 1981; Knorr, 1977; Patton et al., 1977; Weiss & Bucuvalas, 1977), a time we might call the "golden age" of research on evaluation use. More recent work on use has expanded the taxonomy of forms of use, highlighting the notion of process use (e.g., Patton, 1997; Russ-Eft, Atwood, & Egherman, 2002).

A recent resurgence in writings on evaluation use seems to have occurred (e.g., Alkin, 2003; Caracelli, 2000; Cousins, 2003; Henry, 2000; Henry & Rog, 1998; Hofstetter & Alkin, 2003; Kirkhart, 2000; Patton, 1997; Shulha & Cousins, 1997). Nevertheless, as with other aspects of evaluation practice, these contemporary writings on use rarely involve systematic, comparative research (Henry & Mark, 2003), instead being predominantly conceptual or theoretical. Still, some research on use continues to be conducted, probably more than on any other aspect of evaluation practice. Examples include research on the outcomes of evaluability assessments (Rog, 1985), a survey of evaluators on the topic of use (Preskill & Caracelli, 1997), empirical research on use that has cropped up from time to time in related fields such as policy analysis (Greenberg, Mandell, & Onstott, 2000; Shulock, 1999), and post hoc analysis of the processes thought to have led to use or non-use in an actual evaluation (e.g., Russ-Eft et al., 2002). In short, use remains a focus of theory and research on evaluation, as well as a key goal identified by most evaluation practitioners (Preskill & Caracelli, 1997).

Despite its central role in evaluation theory, research, and practice, existing conceptualizations of use include significant gaps and shortcomings. One key gap is that the literature on use has generally failed to give adequate attention to the intrapersonal, interpersonal, and societal change processes through which evaluation findings and process may translate into steps toward social betterment. Imagine trying to translate the literature on use into a general "program theory" of evaluation itself. The literature, especially as summarized in discussions of the different kinds of use, would give limited guidance about the underlying processes (or mechanisms or mediators) that should be included in the program theory of evaluation. A central purpose of this paper is to begin to fill this gap.

Inadequate attention to change processes and interim outcomes is not, however, the only shortcoming of the literature on use. Critics contend that use is inadequate as a motivation

for evaluators, because it fails to provide a moral compass (House, 1995) and falls short of recognizing and promoting the ultimate purpose of evaluation as social betterment (Henry, 2000; Mark, Henry, & Julnes, 2000). A related challenge involves differentiating misuse from use, especially in light of the values conflicts which abound. When disagreements exist about both means and ends of social programs, one person's use could be another person's misuse (Shulha & Cousins, 1997). Yet another concern is that use draws on a conceptual and linguistic base that is too narrow, relative to the broad berth of evaluation influence (Kirkhart, 2000).

In the next section, we begin to show what it would mean for evaluation theory, research, and practice to focus on evaluation influence, to attend to the change processes that may mediate the effects of evaluation findings and practices, and to specify the expected interim outcomes that lie between evaluation activities and the ultimate goals which justify conducting evaluations. Thus, in the following sections, we lay out a theoretical groundwork specifying the mechanisms through which an evaluation can trigger its outcomes, as well as the short- and long-term outcomes of interest. We do this by:

1. Defining the consequences of evaluation and a set of terms that can be used to sort out various types of outcomes and impacts that are associated with evaluation.
2. Providing an alternative framework to specify and organize the ways that evaluation can be expected to influence changes in attitudes and actions at the individual, interpersonal, and collective levels.
3. Summarizing research-based examples from evaluation and other fields for greater rigor and specificity in designing research on the consequences of evaluation or planning for greater impacts of an evaluation.

Finally, we join Kirkhart (2000) in advocating for moving the field of evaluation beyond use and toward a focus on evaluation influence.

TOWARD A PROGRAM THEORY OF EVALUATION: DEFINING THE RELEVANT CONSEQUENCES OF EVALUATION

When an evaluation is conducted, many things happen. Some of the consequences are direct, but incidental to the primary social purposes of the evaluation. Evaluators are employed, data are collected, paychecks are drawn, and people meet who would not meet otherwise. Other consequences are more subtle, often indirect, and can take longer to occur. Novel findings lead to insights about programs or participants. Things are learned. Actions that affect service delivery are taken that otherwise would not be. Theories of evaluation influence cannot and should not attempt to direct attention to all of the possible consequences of evaluation. A theory of evaluation influence should include a way of directing attention to some possible effects (e.g., actions affecting service delivery) but not others (e.g., evaluators receiving paychecks). We propose that a theory of evaluation influence should focus on the *subset of evaluation consequences that could plausibly lead toward or away from social betterment*. Social betterment refers to the improvement of social conditions. Put differently, in the context of democratic society, social betterment refers to bringing about a state that would be considered as better than the state that existed before, as judged though deliberation and by public opinion (Henry, 2000; Henry & Julnes, 1998; Mark et al., 2000).

Viewed in this way, evaluation can be understood as an activity that is analogous to an intervention or a program. Evaluation represents a change in or contribution to ongoing processes

that produce consequences, good, bad, neutral, mixed, or indeterminate. From the near infinite variety of changes that can arise as a result of an evaluation, we separate the consequences of interest from more incidental consequences and rely on the term "evaluation outcomes" to refer to the consequences of interest. Evaluation outcomes of interest are those that are plausibly linked to the ultimate purposes of evaluation, social betterment. For example, evaluation outcomes that fall within the framework presented below include changes in an individual's attitudes about a program, raising an issue onto the public policy agenda, or inclusion of evaluation findings in a speech to justify a program proposal. For each of these outcomes, a plausible case can be made that it could serve as a step toward or away from social betterment. Falling outside the framework are consequences such as income, employment status, or professional prestige that evaluators accrue from conducting an evaluation, and personal relationships that form that do not affect social change, for example. Falling between these two are consequences, such as institutionalizing evaluation or increasing evaluation capacity, which may be considered outcomes of interest but which should be linked to the actual evaluations conducted, adding a step or steps to the "pathway" to betterment. Admittedly, the evaluation outcomes of interest, at least the intermediate and short-term ones, can be subtle and sometimes indirect. The more distal outcomes should be salient and clear, corresponding to the substantial purposes for which evaluation is conducted and which justify use of public resources for evaluation.

To take the analogy of an evaluation to an intervention a bit further, one way to think of this paper is as an attempt to develop an evaluation theory that is similar to a program theory. A program theory is a "small-t" theory that connects program activities and outputs with outcomes. Program theory emphasizes tracing the processes that provide a plausible link between the program activities, on the one hand, and the ultimate goal(s) for which the program was designed and implemented, on the other. The types of program theories to which we refer include underlying processes (also known as mediators or mechanisms) that are external to the program processes and interim outcomes. Analogously, a theory of evaluation influence can be developed to connect evaluation activities and outputs with evaluation outcomes, identifying potential linkages between immediate, intermediate, and long-range outcomes that are indicators or precursors of the specific long-range social goals that motivate evaluation. Such a theory can serve as a tool for planning evaluation activities (e.g., to guide evaluators' actions about how to achieve influence) or as a guide for research on evaluation influence.

Theories of influence can be local and individualized to a particular program evaluation and context, or they can be general and designed to apply across a wide range of contexts. In the next section of the paper, we do not offer a local theory of evaluation influence, but instead provide a more general framework that defines and organizes processes and outcomes that are likely to operate in a variety of evaluation contexts. These concepts can be adapted, however, to develop more specific, local theories of evaluation influence for a given evaluation. Not all evaluations will reach or attempt to reach their goals through similar pathways, and few if any will involve the entire array of change processes that exist. A more encompassing and general theory of evaluation influence, as sketched out in the next section, will not only fill a gap in existing models of use and provide a guide to future research on evaluation influence. In addition, such a framework can provide the building blocks for local theories of influence.

Fortunately, the framework need not be developed de novo. Social science provides both theories of change that are relevant and research on specific outcomes that are similar to those that can be expected to appear in chains of outcomes through which evaluation could lead to social betterment. Just as it is common in program theory development to draw on extant social science theory and research, the framework in the next section draws on empirical and

theoretical work in the social and behavioral sciences. Existing research literatures specify change processes that can be triggered by evaluation, and point the way to constructing indicators of the processes and their outcomes. By drawing on existing, well-validated social theories and empirical research, the framework should result in a more powerful influence theory, with more plausible linkages between evaluation activities and desired forms of influence The considerable validation that these processes have received also suggests that they are likely to operate in a wide range of evaluation contexts.

In the next section, then, we present a framework which distinguishes between mechanisms and interim outcomes of evaluation that take place at the individual, interpersonal, and collective levels. The different levels and types of outcomes have generally been left confounded or undifferentiated in much of the prior work on use. The change processes and outcomes in the framework can be organized into outcome chains that correspond to the ways that evaluation's consequences are linked together. Such pathways are illustrated later in the paper. While additional change processes and associated evaluation outcomes can and, we hope, will be identified, those provided in the next section can serve as starting points for research on evaluation influence and planning for influence as well as a reference for future developments.

TOWARD A THEORY OF EVALUATION INFLUENCE: MULTIPLE PATHWAYS AND MECHANISMS

In this section, we present a framework for developing a better understanding of evaluation influence. As part of that framework, we classify change processes and outcomes that evaluations can influence in terms of three levels: individual, interpersonal, and collective.[2] The levels indicate the locus of the change process. For instance, an individual-level process may take place within a single person, while an interpersonal process takes place in an exchange between two or more people, and a collective process might occur in an organization or institution. As examples, reading about an evaluation's findings could lead to a change: in Susan's thinking (individual), in the interactions between Susan and Juan as she tries to persuade him (interpersonal) about something related to the evaluation, or in a legislature decision reflected in a majority vote (collective).[3]

After describing various forms of evaluation influence at these three levels, we present another part of the proposed framework by examining the way one process or outcome can trigger another, sometimes moving from one level of analysis to another (e.g., from the individual to the interpersonal to the collective). In other words, we discuss the multiple alternative pathways through which evaluations can achieve their ultimate effects.

Three Levels of Evaluation Influence

These three levels correspond to different kinds of change processes that have been studied in the social and behavioral sciences. The first level, individual, refers to those cases when evaluation processes or findings directly cause some *change in the thoughts or actions of one or more individuals*, or, more specifically, when the change process or differences in question predominantly takes place within an individual. For example, evaluators often expect that exposure to evaluation findings will change people's beliefs and opinions. Alternatively, the notion of process use suggests that an individual may learn something or change her beliefs based on participation in an evaluation (note that the difference between evaluation

findings and evaluation process as the stimulator of change corresponds to what Kirkhart, 2000, labels the *source* of influence, not to the change process itself). Foreshadowing our later discussion of evaluation pathways, it is important to keep in mind that one element in the framework we present may stimulate another, either within or across levels of analysis. For example, evaluators may often assume that change in individual attitudes will trigger changes in individual behaviors, or lead to change at the collective level.

The second level in our framework, that is, the interpersonal, refers to a *change brought about in interactions between individuals*, or more precisely, to a process or outcome that predominantly takes place within interactions among individuals. For example, evaluation findings may constitute an authoritative source that a speaker relies upon to try to change the attitudes and behaviors of others. In program and evaluation contexts, the process of persuasion, for example, is most often interpersonal, with one (or more) individuals striving to influence others. The third level, the collective, refers to the direct or indirect *influence of evaluation on the decisions and practices of organizations, whether public or private*. More precisely, the collective level of analysis is involved when a change process or outcome operates predominantly within some aggregate social organizational body. For example, a formal policy change might be influenced by evaluation findings.

To orient the reader to the taxonomy, these three levels, and specific forms of influence within each level, are summarized in Figure 1. Under each of the levels, we have listed between four and six specific types of change processes or outcomes that could plausibly be triggered by an evaluation. Each of these effects has been drawn from a research literature in the social and behavioral sciences that has examined change processes. Indeed, most of the forms of influence we discuss have deep roots in the social sciences and have proven fruitful for research and theory on change processes. In specifying the various forms of influence, we have borrowed heavily from research in psychology, political science, organizational behavior, and elsewhere. One of
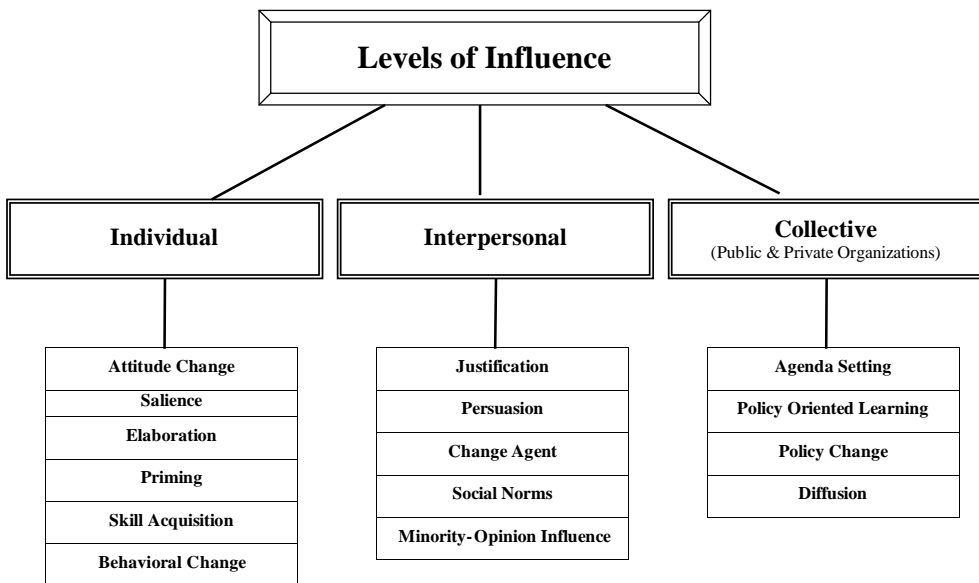

Figure 1.   Mechanisms through which evaluation produces influences.

TABLE 1.
Change Processes for Evaluation Influence at the Individual Level

Individual Mechanisms	*Example*	*References*
Attitude change	Evaluation findings affect state policymakers' opinions about feasibility of implementing program	Greenberg et al. (2000)
Salience	Information about program effects raise the importance of the issue or social problem	Henry and Gordon (2001)
Elaboration	Evaluation report stimulates individuals to think more about the program and their expectations for its outcomes	Petty and Cacioppo (1986)
Priming	Discussion of test scores at beginning of evaluation report makes test scores more important in reader's judgment about program effectiveness	Krosnick and Brannon (1993)
Skill acquisition	New skills, such as collaboration or survey techniques, are learned through participation in evaluation	King (2002)
Behavior change	Teachers begin to mix whole language and phonics instruction after learning that combining the approaches is more effective that either one separately	Patton (1997)

the factors that may have stunted research and theory on evaluation use is that the literature on use has developed terminology that inhibits rather than facilitates the integration of knowledge from other areas that are examining similar change processes.

Several points about Figure 1 are worth noting. First, we do not expect every study of influence to consider all of the processes and outcomes listed in Figure 1, or every practitioner's plan to address all avenues to influence. Instead, Figure 1 offers a menu for researchers, using either preordinant or emergent designs, and practitioners to consider. Second, the list of processes and outcomes in Figure 1, through the associated social science literatures, should add to what is known from research on and conventional wisdom about use with additional tests of factors related to evaluation influence. Third, consideration of the social change processes we have identified could feed into future discussions of evaluation training and evaluator competencies. Fourth, judgments were admittedly involved in selecting the elements listed in Figure 1. However, we believe the framework summarized in the figure could stimulate work that can lead to further development and refinements. Our treatment of the change processes is brief, designed to lead up to our discussion about how these separate outcomes can be linked into chains that constitute pathways through which evaluations realize influence.

Individual Level Influences

We identify six specific mechanisms and measurable outcomes of evaluation at the individual level: direction or valence of an attitude, salience, elaboration, priming, skill acquisition, and behavioral change. The first four mechanisms affect individual attitudes about a program, its importance, or its underlying assumptions. Behavioral change connotes changes in what individuals do, in terms of the delivery of or participation in a program or process. Skills involve both cognition and action. Each of these outcomes that could indicate the influence of an evaluation has fruitfully been investigated elsewhere in the social and behavioral sciences. In Table 1, each of the individual-level attitudinal and behavioral forms of influence is listed

and characterized in two ways:

1. a brief description of an example in which it appears that outcome actually was or plausible would be affected; and
2. references to related research.

Given the number of specific change processes and readily available literature, we describe and illustrate them only briefly here.

Direction or valence of an attitude change.   Individuals may hold positive or negative beliefs and attitudes about a program or one or more of its underlying assumptions. Changing an individual's attitude from positive to negative or vice versa is one end-state that most evaluators would cite if asked to describe how an individual would be likely to "use" an evaluation. Attitude change may encompass much of what most evaluators would characterize as "enlightenment," but we believe enlightenment may include more, such as awareness that other attitudes could be held, and is more ambiguous than changing the direction of an attitude. A particularly relevant example of attitude change comes from a recent study of the impacts of welfare reform evaluations. Greenberg et al. (2000) report that state policymakers shifted their attitudes about work-oriented reforms when they read evaluations that established the feasibility of implementing reforms of this type and described the impact of these reforms in other states. In this case, changes in individual's attitudes from negative to positive preceded a collective action, policy change, a possibility we return to in the later discussion of pathways of influence.

Salience.   Salience refers to the importance of an issue as judged by an individual. An example of salience as an outcome of evaluation is provided by recently released evaluations of the effects of welfare reform on children. Several high visibility evaluations, released in advance of the welfare reform reauthorization, were intended to raise the importance in people's minds of children's well-being as a consequence of welfare reform (Fuller, Kagan, Caspary, & Gauthier, 2002; Zaslow et al., 2002). In other words, the evaluations were commissioned to raise the long-term salience of this outcome, so that it would become a concern when citizens and legislators think about welfare reform, specifically in the context of reauthorization of the act that created Temporary Assistance to Needy Families (TANF).

Elaboration.   Elaboration refers to an important aspect of the opinion formation process, specifically, how much an individual actually thinks about an issue. Elaboration is the key element of "central processing" of information about an issue. It stands in contrast to "peripheral" or "heuristic" processing in which an individual relies on other cues, such as who is supporting an issue, to form their opinions about an issue (Petty & Cacioppo, 1986). Elaboration can be stimulated by participating in an evaluation or by thoughtful reading of its findings. Elaboration is a process that can lead to attitude change, to more stable attitudes, to more (or less) certainty in one's attitudes, and to other outcomes that indicate influence.

Priming.   Generally speaking, priming describes the phenomenon by which a concept temporarily becomes cognitively accessible. Because a primed concept is at the "top of the mind," it can influence people's judgments about individuals (Higgins, Rholes, & Jones, 1977; Sinclair, Mark, & Shotland, 1987), institutions or office holders (Krosnick & Brannon, 1993; Krosnick & Kinder, 1990), events (Nelson, Clawson, & Oxley, 1997), or, we presume, policies

or programs or even evaluations. Priming can affect the very *way* one thinks about an issue. In general, priming a concept makes that concept and related ones more influential in the judgment process (and these effects on judgments may persist even if the priming itself does not; Sinclair et al., 1987). An evaluation or evaluation report can trigger priming, for example, by temporarily highlighting one kind of outcome rather than others in the findings. Or priming could arise from the environment, for instance, from the morning newscast. Whether built into the evaluation report or stemming from the environment, priming might influence judgments about programs and policies, relative to the judgments that would be made in the absence of the primes. These judgments affected by priming could then alter the attitudes and subsequent actions of those who have been primed.

Skill acquisition. Skill acquisition refers to the development of new abilities. The source of the skills can be participation in an evaluation. For instance, evaluation participants may learn about survey methodology by designing, conducting, and analyzing surveys (King, 2002). Certain kinds of skill acquisition, such as participants' ability to work collaboratively or developing persuasive arguments based in data from systematic research, could later lead to other outcomes. Another example comes from an evaluation in which teachers learn to rigorously assess portfolios of students' work as a part of collecting data. Teachers who learn these skills may then apply them in assessing their own students' work and subsequently individualize instruction based on the assessments.

Behavior change. Behavior change represents direct instrumental use at the individual level. As an example, individual classroom teachers may change their approach to reading instruction after they learn about an evaluation which shows that mixed reading methods produce more consistent results across diverse groups of children than either "whole language" or phonics. Behavior change at the individual level appears to be an important outcome in some evaluation theories, including some participatory and empowerment approaches (e.g., Fetterman, 2001).

Interpersonal Level Influences

We have borrowed from other active research areas in identifying five different forms of interpersonal behavior that may be affected by evaluation: justification, persuasion, change agent, social norms, and minority-opinion influence. These represent a variety of processes through which one or more individuals can undertake to influence the attitudes or actions of others. In other words, several different forms of influence exist at the interpersonal level, and each can play a mediating role in a causal chain of evaluation influence, often following an individual level change of the person attempting to influence others (see Table 2).

Justification. Justification refers to the use of evaluation findings to support a previously held position. In some previous discussions, justification has been darkly tainted (Knorr, 1977). We disagree. Political scientists have established the necessity of presenting a convincing case to obtain approval of an initiative in a variety of political processes (Majone, 1989). Justification should only receive a pejorative label when it involves invoking inaccurate or incomplete statements about the findings of an evaluation. For example, it is hard to criticize a proponent of high quality child care for taking an evaluation that finds that high quality child care leads to positive child outcomes and publicizing the findings to justify her proposal. A recent example

TABLE 2.
Change Processes for Evaluation Influence at the Interpersonal Level

Interpersonal Mechanisms	Example	References
Persuasion	Communication of results through credible organization persuades policymakers that program should be supported	Greenberg et al. (2000)
Justification	Results showing quality child care enhances children's development are used to justify higher quality standards for child care licenses	Shulock (1999)
Change agent	Results influence an individual to work for organizational change	Weick and Quinn (1999)
Minority-opinion influence	Opinion minorities use findings to counter widely held attitudes	Glazer (1994)
Social norms	Evaluation changes social norms about telecommuting and new norms increase telecommuting and reduce driving	Ajzen (2001), Henry and Gordon (2003)

of research on justification comes from the field of policy analysis. Shulock (1999) counted citations of policy analyses and evaluations in Congressional Committee reports as a measure of the attempts by these committees to justify their jurisdiction over specific pieces of legislation. She found that when the legislation involved issues of high public interest and when committees were competing with each other for jurisdiction over the legislation, there were more extensive citations of the evaluations and policy analyses. Justification is an important outcome for an evaluation, and it may mediate other outcomes. In many cases justification is an initial step in a process that is designed to persuade others, the next outcome that we will consider.

Persuasion. Persuasion refers generally to the attempts of one party to change the attitudes of another, usually through direct communication. An example of persuasion can be drawn from the research of Greenberg et al. (2000) discussed earlier. National networking organizations, such as the American Public Welfare Association and the National Governor's Association, through their choice of speakers at national conferences disseminated persuasive messages about the evaluations of welfare reform. This interpersonal persuasion effort apparently triggered a subsequent reversal in attitudes of state policymakers who had not been directly exposed to the evaluation findings about the feasibility of welfare-to-work programs.

Change agent. A change agent is someone focused on changing organizational policies, programs, or operations. Our focus here is on those cases when evaluation findings or processes lead an individual or individuals to rise to action as a change agent. Evidence from an evaluation may stimulate this reaction, or participating in an evaluation may embolden it. The successful change agents may in turn be persuasive, which may then trigger other kinds of change.

Social norms. Social norms are shared standards about what is an appropriate or acceptable behavior in a specific context and at a particular time. Communication about norms can be subtle. But norms are inherently interpersonal, in that they involve the effects of others on the individual. Perceived norms can become internalized by individuals, thereby influencing

action. Norms can influence behavior even if they are not internalized, if the person thinks others will behave less favorably if he or she violates the norm. As an example, welfare case workers may, through an evaluation, hear about other case workers' actions regarding ensuring that all welfare applicants are fully informed about the range of possible benefits for which they may be eligible; from this, the case workers may perceive a social norm and this may influence their behavior.

Minority-opinion influence. Minority-opinion influence is a specific category of influence that has been discussed in the psychological literature. The term minority refers here to those in the minority with respect to attitudes (not demographics), and often is applied to those who advocate positions outside of the mainstream. Research in this area, which began in reaction to the voluminous literature on conformity, was intended to demonstrate that, under the appropriate conditions, those holding minority views could overturn the majority. Early research findings suggest that minorities are most likely to be effective in bringing about change when they are persistent, when they have a strong argument with a moral component, and when they have charismatic leadership (Moscovici, Lage, & Naffrechoux, 1969). Experiences of groups such as Mothers Against Drunk Drivers (MADD), opponents of public smoking, and children's rights advocates, indicate how information can be marshaled in support of minority-opinion influence (Glazer, 1994). Evaluations can and often are the source of this information.

Collective Action by Public and Private Organizations

In addition to the individual and interpersonal level change processes, the literature on evaluation outcomes is replete with expected changes at the collective level. Whether from the early literature on direct instrumental use of evaluation findings or more recent work that catalogues experiences with "organizational learning," evaluation theorists have often pointed to the importance of organizational level change. We employ the term collective rather than organizational because the outcomes can be measured across individuals or across multiple organizations or at an institutional level that may not be organizational, per se. We focus attention on four outcomes that represent collective level changes: agenda setting, policy-oriented learning, policy change, and diffusion. Each of these could be influenced by evaluation, perhaps at a later (or even final) step in a complex pathway of influence (see Table 3).

Agenda setting. Agenda setting refers to an issue moving onto the media agenda, the public agenda, or the government agenda (Zhu, 1992). Students of communication and policy analysis and advocates of particular public policies have long considered media reports of research results, including evaluation findings, as crucial step leading to policy change. A substantial body of research demonstrates that media coverage influences public opinion about the importance of the issue (McCombs & Zhu, 1995; Neuman, 1990), and that issues that are more important to the public are likely to be acted upon (Monroe, 1998; Page & Shapiro, 1992). More and more we observe evaluations being reported in the media (e.g., Reynolds, Temple, Roberston, & Mann, 2001) in an attempt to trigger a chain of related actions that lead to evidence-based policy change. Given the limited carrying capacity of the media, having an evaluation related issue reported in the news is a non-trivial outcome. If the reporting elevates the public's attitude about the importance of the issue to the current list of what people are most concerned about, there is a much greater likelihood that the issue will receive attention by

TABLE 3.
Change Processes for Evaluation Influence on Collective Actions

Collective Action Mechanisms (Public or Private Organizations)	Example	References
Agenda setting	Media reports of evaluation findings increase public opinion about the importance of providing preschool	Henry (2002) and McCombs and Zhu (1995)
Policy-oriented learning	Advocacy coalition modifies policy recommendation based on evidence of success of several program alternatives	Sabatier and Jenkins-Smith (1993)
Policy change	Evidence of growing social problem influences the adoption of a major policy reform	Mintrom (1997)
Diffusion	Evidence of a policy's success in another jurisdiction influences the adoption of similar policy elsewhere	Berry and Berry (1990)

policymakers. Thus, agenda setting is an important outcome that may be a precursor to other forms of collective change.

Policy-oriented learning. Policy-oriented learning refers to "relatively enduring alterations of thought or behavioral intentions that result from experience and/or new information and that are concerned with the attainment or revision of policy objectives" (Sabatier & Jenkins-Smith, 1998, p. 123). This type of learning is a counterpart of the change in individual attitudes at the collective level, involving coalitions of advocates who are actively seeking to have their policy preferences adopted. A single evaluation may be unlikely to change pre-existing beliefs because "actors tend to accept information confirming existing beliefs and to screen out dissonant information" (Sabatier & Jenkins-Smith, 1998, p. 145). However, policy-oriented beliefs are different from individual attitudes in several important ways. For example, because these are beliefs held by a coalition, they are subject not only to change by turnover in the leadership of a coalition, by change in the membership of the coalition, when coalitions become dominant within an institution as well as by attitude change at the individual level. Thus, new leaders with different policy core beliefs may change the "take away" points or lessons learned from evidence about the effectiveness (or ineffectiveness) of certain policies or regulations.

Policy change. Policy change is the type of influence that evaluations were originally expected to bring about (e.g., Suchman, 1967). Policy change refers here either to major or to minor, more incremental modifications of policies (Sabatier & Jenkins-Smith, 1993) within the environment where the evaluation occurred. Mintrom (1997) provides an example illustrating the effect of a particular kind of evaluative information, state test score results, on policy change, specifically school reform legislation. Perhaps not surprisingly, the influence of evidence on policy change was complex and somewhat nuanced. Mintrom (1997) found that state test score results affected the actual *adoption* of school reform legislation, but not the *consideration* of such legislation. In addition, major policy changes were more likely to occur within states where test results indicated that the performance of the schools statewide was declining. Of course, other factors also influenced the adoption of school reforms, but Mintrom's research

suggests an important hypothesis with respect to the valence of evaluation findings. Negative findings, or findings that show a growing problem, may be more influential for policy change than are other findings. On the other hand, we might expect that positive findings are more likely to influence policy consideration in other jurisdictions (diffusion) that would like to capture similar benefits. In short, the nature of the findings from an evaluation may trigger different pathways for influence.

   Diffusion.   Diffusion refers to the adoption of the policy, program, or practice being evaluated in a jurisdiction outside of the one in which the evaluation was conducted. Greenberg et al. (2000) report some diffusion effects from the evaluations of a set of major welfare to work programs, GAIN (CA), CAP (NY), and Project Independence (FL). For example, representatives from 30 different states "recalled that GAIN was specifically mentioned at the time that their state was considering moving their JOBS program toward a labor force attachment model" (2000, p. 375).
   Both policy change and diffusion are commonly categorized in taxonomies of use as direct instrumental use (but see Hofstetter & Alkin, 2003, for a more restrictive definition of use that seems to exclude diffusion). Both policy change and diffusion are important and tangible forms of evaluation influence at the collective level. However, as suggested by earlier discussion of the other levels of influence, these collective changes are likely to be preceded by individual and interpersonal changes. In other words, these collective actions may come at or near the end of a multiple-step pathway of influence.

Multiple Pathways of Evaluation Influence

   In practice, specific outcomes (such as change in an individual's attitudes and interpersonal persuasion) will often be connected with each other like the links in a causal chain. Because the influence of a single evaluation can transpire through numerous outcome chains, there are multiple possible pathways of influence. To take but one example, a pathway of influence might begin with a person who learns from a multi-site evaluation that the local program was not producing the expected outcomes and, as a result, changes her attitudes about the local program. This individual-level attitudinal influence might then cause the person to take on the role of change agent. In turn, this specific type of interpersonal behavior might, after some time, lead to the adoption of a more successful program model from another site in the evaluation, that is, to diffusion. But it could be that a window of opportunity was opened in the policy agenda by other, unrelated events without which the policy would not have gained consideration. Evaluation is a potential influence at all of these levels, but the goal of achieving social betterment may be thwarted by other elements in the environment that counteract the influence of the evaluation or, more benignly, fail to create a hospitable environment for the evaluation's influence to move further along the desired chain.
   We could posit a case in which the findings from an evaluation arm a group whose attitudes are initially in the minority with respect to an existing social problem and the group publicizes the results through the media attention and in meetings with individual legislators, which subsequently raise the social problem's perceived importance (i.e., agenda setting). Note that this particular pathway does not begin with attitude change. Rather, the causal chain begins with the evaluation findings that energize the opinion minority group and give it a tool to use in its effort to raise their concerns on the public agenda (by publicizing the evaluation findings in an attempt to raise public concerns). When the issue rises on the public's agenda, individual

members of the public may begin to think about the issue more and they begin to elaborate their opinions, perhaps resulting in attitude change about the desirability of public action on the issue. Finally, a collective action occurs, in the form of a change in policy, which some research indicates can in fact result from agenda setting and attitude elaboration (Monroe, 1998; Page & Shapiro, 1992).

Borrowing an example slightly outside of the evaluation sphere, this pathway roughly corresponds to the way groups such as Mothers Against Drunk Driving have exerted influence to change policy. Beginning with a minority position (about enforcement activities and criminal penalties for drunk driving), MADD members armed themselves with statistics (as well as heart-rending case examples of fatalities). Their efforts as a minority opinion group helped put the issue of drunk driving on the forefront of public opinion. Presumably because of the thinking this stimulated about the issue (i.e., elaboration), attitude change occurred. Whereas the public once considered drinking and driving a private issue with the risk accruing to the individual, opinions changed when the behavior was seen as placing others at risk. Policy change followed, in the form of tougher laws about drinking and driving.

Notice that this account provides considerably more detail than simply saying that "the statistics about alcohol-related traffic fatalities were used instrumentally." This kind of analysis also can help provide insight into why a particular evaluation (or other change effort, more generally) fails to effect some kind of change. Imagine, for example, that shortly after the efforts of MADD to put drunk driving on the public agenda, the U.S. was involved in a war. This could have taken all the oxygen out of news coverage and public discourse about drunk driving. Would that have been the fault of the evaluator or represent a shortcoming of the evaluation? We think not, though it might seem that way from the normative perspective on use. Or imagine that the issue climbed for a time onto the media agenda, but people did not elaborate their thoughts about it or change their attitudes, as appears to have happened with the issue of abuse of the elderly in federally funded, home health care programs (Cook et al., 1983). In short, thinking about the pathways of evaluation influence allows an opportunity to understand and study when and why some end-state use occurs and when and why it does not. Modeling the causal chain of a planned change process, like laying out a program theory for a new program, can also allow evaluators and other participants in the evaluation process to assess the plausibility of the plan in the first place.

In the previous example, we presented a chain that originates from the evaluation findings, rather than evaluation process (Patton, 1997), although both are viable starting points for evaluation influence. The influence of the evaluation process is initially more limited in scope than is the potential initial influence of evaluation findings because access to the process is often more limited than access to evaluation reports. That is, a chain triggered by the evaluation process must necessarily begin with participants in the evaluation, rather than with the audience who could potentially read or hear about the evaluation findings. Nevertheless, a process-triggered chain can still be potent, especially when key opinion leaders or decision-makers are included. For example, teachers in a participatory evaluation might learn to apply new assessment methods as a result of their decision to assess portfolios of student work. The teachers could decide to apply their new skills in assessing the work of their students more frequently and change their instructional practices as a result. If these practices result in greater student success, these assessment practices could become a part of the school's social norms and be adopted by others, especially beginning teachers entering the school, with subtle normative pressure encouraging these same practices.

Conclusions About Pathways

There are many, many other causal pathways that could be put together using the building blocks of influence processes listed in Figure 1. Some pathways are probably fairly common in practice. Others may never happen. Some pathways may be far more likely than others to lead to some desired end state, such as policy change. Of course, we will never be clear about which do and which do not occur unless we move beyond use to a finer-grained model of evaluation influence, one which includes explicit linkages between the multiple outcomes along the pathway of influence. This kind of orientation is also needed for a fuller understanding of the cases that would now be labeled as instances of "non-use." We need a more detailed framework to guide the question: Where did the train of influence derail and why? We also need to add contextual factors to the "pathways" to better understand the outcomes for which an evaluator might be held accountable and those that are part of a larger theory of change but which are beyond an evaluator's scope of influence.

The theory of evaluation influence presented here has several implications for future research and development of evaluation influence. Let us consider but four of them.

(1) It may be relatively easy to identify the expected first link in the chain of influence for a given evaluation. It may also be easy to identify the change processes that are thought to be common across many evaluations. For example, many common pathways are likely to begin with individual attitudinal effects. Fortunately, these are also for the most part the easiest outcomes to measure. It might contribute substantially to an evidence-based practice if evaluators more routinely assessed (at least some of) these attitudinal outcomes. Do they differ before and after an evaluation or not? Data like this collected from multiple studies could allow interesting comparisons across evaluations, as well as comparisons before and after a particular evaluation.

(2) Just as in program theory-driven evaluation, if there are no effects on the initial steps in an expected influence pathway, it puts the longer-term outcomes in doubt. Perhaps, for example, instrumental use is so uncommon precisely because it is premised on influence pathways that begin with attitude change, and because single evaluations rarely stimulate attitude change. We will never know for sure whether this is in fact the case, until we start taking the study of influence pathways seriously. We may find, for example, that it is actually turnover in leadership among elite members of advocacy coalitions that affects policy beliefs more than attitudinal change of those in leadership positions. As a first step, we begin to make our assumptions explicit when we lay out the pathways that are expected. For example, in a recent study of the effectiveness of an information campaign, the program planners placed awareness in the causal chain between the campaign activities and changes in behaviors. In fact, greater levels of awareness were associated with less behavioral change (Henry & Gordon, 2003). The evaluation, which tested the linkages between the outcomes in the pathway, caused the chain to be reformulated, still including awareness as an outcome but not expecting a relationship with behavioral changes. Until we study influence pathways, we will be hard pressed to figure out how evaluation can lead to more evidence-based social change and how evaluation can be carried out to achieve more influence.

(3) The theory presented here can support a variety of kinds of future research on evaluation use. For example, quantitative multivariate studies could be done collecting data on evaluation context and process and on various measures of influence, to examine

the relative influence of several possible determinants of influence. Qualitative case studies could carefully track the details of influence pathways associated with a particular evaluation or set of evaluations as they emerge. Both conceptual and empirical work could be undertaken to refine or extend the taxonomy of influence processes summarized in Figure 1.

(4) Simply having a new framework or theory may help stimulate new research on the consequences of evaluation. The approach summarized in this paper, we believe, may be especially likely to stimulate research because it should facilitate evaluators' awareness of, interest in, and learning from other areas that are also interested in the influence of information and in individual, interpersonal, and organizational change. The "program theories" are testable in the sense that the indicators of the outcomes are measurable, although they may not necessarily be measured, and the linkages made explicit in the outcome chain are testable. Indicators of outcomes can be specified and measured at appropriate time intervals. Research designs can be worked out to test impacts using causal analysis or to monitor outcomes descriptively over time. Other social science research literatures can be a powerful source of methods, measures, and hypotheses. If interplay increases between evaluators who study evaluation influence and those in other research traditions that study change processes, it is likely to be beneficial for both. The interplay is likely to be most productive where there is substantial overlap of interests in the specific outcome between the two groups. For instance, substantial overlap exists between political scientists who study agenda setting and evaluators interested in the effects of formal evaluation findings on the mass public's perceptions of the importance of an issue addressed by the evaluation. Because of this, progress in the field's understanding of evaluation influences may be somewhat selective and opportunistic, testing mostly theories of interest and current import to both evaluators and those in allied areas studying influence and change process. However, selective progress is presumably better than no progress at all.

FROM USE TO INFLUENCE

Research, theory, and practice related to use would benefit from increased attention to the mechanisms, outcomes, levels of analysis, and pathways described in the previous section. However, even more benefits may accrue if the field of evaluation were to move beyond use to an emphasis on evaluation influence (Kirkhart, 2000). Kirkhart's challenge to the current complacency surrounding the term use arose in a series of conference sessions sponsored by the American Evaluation Association's Topical Interest Group on Evaluation Use, which resulted in a volume entitled, "The Expanding Scope of Evaluation Use" (Caracelli & Preskill, 2000). In this volume, several authors provided insights about ways in which the concept of use has: expanded and developed over the years: moved in subtle and substantial ways beyond instrumental, conceptual and symbolic uses: and become integrated into evaluators' attention to organizational learning. In the lead article, Kirkhart (2000) makes the case that "use" draws on a conceptual and linguistic base that is too narrow. She argues for replacing use with influence: "The term *influence* (the capacity or power of persons or things to produce effects on others by intangible or indirect means) is broader than use, creating a framework with which to examine effects that are multidirectional, incremental, unintentional, and instrumental" (p. 7).

In highlighting three dimensions of influence—intention, source of influence, and time—Kirkhart identifies three aspects of evaluations' influence that could be studied. Two of these, intention and source of influence, come prior to the mechanisms and outcomes presented in this paper and summarized in Figure 1. The other, time, represents the temporal dimension during which the mechanisms operate and the outcomes come to pass. Kirkhart offers a compelling rationale for studying influence in general and these three aspects in particular. The reasons that she gives include: mapping the influence surrounding a particular evaluation, sorting out use and misuse, improving the validity of studies of the consequences of evaluation, and, from a pragmatic standpoint, most importantly, reducing underestimates of the impacts of evaluation (pp. 18–20). We add to Kirhart's argument and framework of evaluation influence by highlighting the many social change processes that can result in influence, as well as the interim outcomes that can mediate longer-term outcomes and the complex pathways of processes and outcomes that can be formed.

More generally, we join Kirkhart in arguing that simply adding to categories and definitions that appear under the umbrella of use is no longer sufficient for the field to understand and thoroughly examine the consequences of evaluation. Thus, we argue that the time has come to respond affirmatively to the question: Should the field of evaluation move beyond the concept of use? The previous sections of this paper contribute to an alternative characterization of the consequences of evaluation and a framework for specifying these effects. However, that framework could be adopted by scholars and practitioners by being integrated within existing models of evaluation use. In this section, we go further, by briefly addressing some of the shortcomings of use and benefits of a focus on evaluation influence, as Kirkhart (2000) has advocated. Thus, we argue that attention to the underlying mechanisms and interim outcomes that mediate the valued outcomes of evaluation should take place within a developing theory of evaluation influence, rather than in the older frameworks of evaluation use.

The Limitations of Use and Benefits of Influence

As evaluation theory has developed, most of the discussions of the consequences or outcomes of evaluation have been organized under the construct of use. Use has a long history in the field of evaluation both as an objective for evaluation and as a fundamental concept for research on and theories of evaluation. To be sure, the concept of use has played an important and often valuable role in evaluation theory and practice. For example, evaluators often invoke the word use casually in conversations or negotiations with sponsors or stakeholders. "Use" is a positive and generally non-threatening way of conveying that an evaluation could and, in many cases, should affect policies, practices, and programs. "Use" is shorthand for the practical reason for doing evaluations. There are also "longhand" versions of use that appear in communications among evaluators. That is, more specific frameworks exist that attempt to capture the different possible outcomes of evaluation, in the form of several conceptually-driven (and overlapping) taxonomies of kinds of use (e.g., Caracelli, 2000; Cousins & Leithwood, 1986; Knorr, 1977; Patton, 1997; Shulha & Cousins, 1997; Weiss, 1979), which emphasize such forms of use as instrumental, conceptual, symbolic, and process use, as well as misuse.

While it continues to perform valuable service as a handy catchall phrase, use has largely avoided rigorous and consistent specification, operationalization, or empirical examination. Researchers have discovered and documented a wide variety of outcomes that evaluation

sometimes spawns. While conceptual innovations continue to arise, such as Patton's process use (1997) and the recognition of multiple levels of use (individual and organizational) (Shulha & Cousins, 1997), contemporary research guided by the concept of use is limited both in terms of quantity and conceptual groundings. Researchers have not developed validated ways to measure many of the constructs of use, for example, of conceptual use. Multiple meanings of the term use, which is sometimes used descriptively and sometimes normatively, create miscommunication. As a result of the limited empirical research base, evaluation plans are more likely to be developed relying on trial and error from individual experience or on shared conventional wisdom, than on evidence about successful plans and designs. Shifting to a theory of influence, rather than use, and attending to underlying influence processes could help stimulate a new generation of research that can better guide practice (Henry & Mark, 2003).

Critics also contend that use is deficient as a motivation because it fails to provide a moral compass (House, 1995; House quoted in Alkin, 1990, p. 104) and because it stops short of recognizing and promoting the ultimate purpose of evaluation as social betterment (Henry, 2000; Mark et al., 2000). Positing that misuse is distinct and different from use and that misuse stems from bad purposes, while use represents a change for the better, has been offered as part of a way of providing a moral footing for use (Patton, 1997). However, in societies where values conflicts abound, it is easy to see that ambiguity cloaks the distinction between use and misuse (Shulha & Cousins, 1997). One person's use could be another person's misuse when disagreements exist about both means and ends of social programs.

One of the counterpoints to Kirkhart's (2000) advocacy of influence is that influence combines evaluation use, which occurs within the program context and general time frame of the evaluation, with knowledge use, which is broader (Alkin, 2003; Hofstetter & Alkin, 2003). According to this position, the distinction between evaluation use and knowledge use is worth preserving (Alkin, 2003). Although appealing, in the sense of not adding additional terminology to evaluation, we disagree with this view, for several reasons. First, the very concept of use connotes intentionality and awareness, and evaluation's influence can be indirect and distant, mediated by complex pathways of processes and interim outcomes. Second, the indirect, delayed, and distant effects of evaluation need not always be stimulated by knowledge per se, and thus do not necessarily fall into the category of knowledge use. For example, participation in the evaluation process can eventually have influence in a setting other than where the evaluation took place, as when staff who took part in an evaluation subsequently move to another location and exert social influence. Third, if the distinction between evaluation use and knowledge use is maintained, it results either in fruitless discussion about where the boundary lies between the two, or in common blurring of the lines between the two (e.g., is enlightenment evaluation use or knowledge use?). The concept of evaluation influence removes the need to either reify or blur the porous line between evaluation use and knowledge use. Fourth, attention to evaluation influence should invigorate conceptual and empirical work on the outcomes of evaluation, by providing a stronger connection between evaluation influence and other literatures on influence. If the shift to evaluation influence, plus the framework presented in this paper, helps stimulate empirical evidence about what kinds of evaluation worked in which circumstances to produce which specific outcomes, this would greatly reduce the field's susceptibility to fads and to debates about evaluation approaches that are more about ideology than effectiveness of evaluation (Henry & Mark, 2003; Mark, 2001, 2003).

CONCLUSIONS

The terms "use" and "utilization" will always be handy, informal ways for referring to generally positive and desired consequences of evaluation. Because it is both concise and handy, the term "use" will continue to be useful for some purposes. But for other purposes, "use" may have become overly laden with multiple meanings, and specific forms of use may be too imprecise or too focused on end-states to fruitfully guide research and practice. For example, focusing on policy change alone (or any end-state instrumental use) too readily allows evaluators and those studying evaluation's effects to omit several types of influence that occur upstream. On the other hand, the idea of enlightenment or conceptual use does not provide reasonable specificity to guide measurement of evaluation's outcomes or to insure that casual observers or researchers are talking about the same outcomes. In addition, existing taxonomies of evaluation use are not clear as to the level of analysis. Is it instrumental use if an individual teacher changes his or her instructional practices after hearing of an evaluation, just as it is if a state legislature funds a new educational program? The pathways to these two end-state outcomes are likely to be very different (Weiss, 1998). Theories of use do not help much in differentiating between these two cases or between the very different pathways to them. The framework for studying evaluation influence offered in this paper can clearly distinguish between the two, and can draw out, step by step, the influence pathway to each. By combining change processes into causal pathways, we have assembled a more powerful way to formulate working hypotheses about evaluation outcomes. These working hypotheses can serve two important purposes: (1) to guide research on the influence of evaluation; and (2) to think through a plan for maximizing the influence of a specific evaluation. Attention to change processes could simply be a way to enhance future work on use. The results could fill a gap in the "program theory" of evaluation. However, given the baggage attached to the concept of use and its shortcomings, combined with the benefits of focusing on evaluation influence, we think it is time to move beyond use and continue to develop theories of evaluation influence.

ACKNOWLEDGMENTS

NOTES

1.   The term "use" is often treated interchangeably with the term "utilization" (Weiss, 1988). These two terms have been applied in nearly synonymous ways in the evaluation literature and we do not here attempt to distinguish between the two.

2.   One could create more categories in the theoretical framework. For example, it would be possible to differentiate four levels of analysis, the individual, dyadic, small group, and collective/public. Alternatively, one could try to specify both attitudinal and action categories for each level of analysis. At

this point, we believe that finer distinctions would not be productive, though we see this as an issue for further empirical and conceptual work in the future.

   3.   There might be some debate about the placement of some elements of the framework within the three levels of analysis, but for the most part we believe this is because similar outcome measures may be used across the levels. For example, persuasion at the interpersonal level can be measured by comparing attitudes before and after the interpersonal interaction, and the same measurements could be taken before and after participating in an evaluation as an indicator of change at the individual level. For a complementary and more detailed system for classifying the elements in the present framework, see Mark and Henry (in press).

REFERENCES

Ajzen, I. (2001). Nature and operation of attitudes. In S. T. Fiske, D. L. Schacter, & C. Zahn-Wexler (Eds.), *Annual Review of Psychology*, *52*, 27–58.

Alkin, M. C. (Ed.). (1990). *Debates on evaluation*. Newbury Park, CA: Sage.

Alkin, M. C. (2003). Introduction (to section on evaluation use). In T. Kelligan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 189–196). Dordrecht: Kluwer Academic Publishers.

Alkin, M. C., Daillak, R., & White, P. (1979). *Using evaluations: Does evaluation make a difference?* Beverly Hills, CA: Sage.

Berry, F. S., & Berry, W. D. (1990). State lottery adoptions as policy innovations: An event history analysis. *American Political Science Review*, *84*, 395–415.

Caplan, N. (1977). A minimal set of condition necessary for the utilization of social science knowledge in the policy formulation at the national level. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 183–197). Lexington, MA: Lexington Books.

Caracelli, V. J. (2000). Evaluation use at the threshold of the twenty-first century. In V. J. Caracelli & H. Preskill (Eds.), *The expanding scope of evaluation use. New Directions for Evaluation*, No. 88 (pp. 99–111). San Francisco, CA: Jossey-Bass.

Caracelli, V., & Preskill, H. (eds). *The expanding scope of evaluation use. New Directions for Evaluation*, No. 88, San Francisco: Jossey-Bass.

Cook, F. L., et al. (1983). Media and agenda setting: Effects on the public, interest group leaders, policy makers, and policy. *Public Opinion Quarterly*, *47*, 16–35.

Cousins, J. B. (2003). Utilization effects of participatory evaluation. In T. Kelligan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 245–266). Dordrecht: Kluwer Academic Publishers.

Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, *56*, 331–364.

Fetterman, D. M. (2001). *Foundations of empowerment evaluation*. Thousand Oaks, CA: Sage.

Fuller, B., Kagan, S. L., Caspary, G. L., & Gauthier, C. A. (2002). Welfare reform and child care options for low-income families. *The Future of Children*, *12*, 97–119.

Glazer, N. (1994). How social problems are born. *The Public Interest*, *115*, 31–44.

Greenberg, D., Mandell, M., & Onstott, M. (2000). The dissemination and utilization of welfare-to-work experiments in state policymaking. *Journal of Policy Analysis* & *Management*, *19*, 367–382.

Henry, G. T. (2002). Transparency, stakeholder involvement, and explanation in contemporary evaluation: A review essay stimulated by Success in Early Intervention: The Chicago Child-Parent Centers by Arthur J. Reynolds. *American Journal of Evaluation*, *23*, 235–244.

Henry, G. T. (2000). Why not use? In V. J. Caracelli & H. Preskill (Eds.), *The expanding scope of evaluation use. New Directions for Evaluation*, No. 88, (pp. 85–98). San Francisco, CA: Jossey-Bass.

Henry, G. T., & Gordon, C. S. (2001). Tracking issue attention: Specifying the dynamics of the public agenda. *Public Opinion Quarterly*, *65*, 157–177.

Henry, G. T., & Gordon, C. S. (2003). Driving less for better air: Impacts of a public information campaign. *Journal of Policy Analysis and Management*, *22*, 45–63.

Henry, G. T., & Julnes, G. (1998). Values and realist evaluation. In G. Henry, G. Julnos, & M. M. Mark (Eds.), *Realist evaluation. New Directions for Evaluation*, No. 78 (pp. 53–72). San Francisco, CA: Jossey-Bass.

Henry, G. T., & Mark, M. M. (2003). Toward an agenda for research on evaluation. In C. A. Christie (Ed.), *The practice–theory relationship in evaluation. New Directions for Evaluation*, No. 97 (pp. 69–80). San Francisco, CA: Jossey-Bass.

Henry, G. T., & Rog, D. J. (1998). A realist theory and analysis of utilization. In G. Henry, G. Julnes, & M. M. Mark (Eds.), *Realist evaluation. New Directions for Evaluation*, No. 78 (pp. 89–102). San Francisco, CA: Jossey-Bass.

Higgins, E. T., Rholes, W., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, *13*, 141–154.

Hofstetter, C. H., & Alkin, M. C. (2003). Evaluation use revisited. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 197–222). Dordrecht: Kluwer Academic Publishers.

House, E. (1995). Principled evaluation: A critique of the AEA guiding principles. *Guiding principles for evaluators. New Directions for Evaluation*, No. 66 (pp. 27–34). San Francisco, CA: Jossey-Bass.

King, J. A. (2002). Building the evaluation capacity of a school district. In D. W. Compton, M. Baizerman, & S. H. Stackdill (Eds.), *The art, craft, and science of evaluation capacity building. New Directions for Evaluation*, No. 93 (pp. 63–80). San Francisco, CA: Jossey-Bass.

King, J., Thompson, R., & Pechman, E. (1981). *Improving evaluation use in local school settings*. Washington, DC: National Institute of Education.

Kirkhart, K. (2000). Reconceptualizing evaluation use: An integrated theory of influence. In V. Caracelli & H. Preskill (Eds.), *The expanding scope of evaluation use. New Directions for Evaluation*, No. 88 (pp. 5–24). San Francisco, CA: Jossey-Bass.

Knorr, K. D. (1977). Policymakers' use of social science knowledge: Symbolic or instrumental? In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 165–182). Lexington, MA: Lexington Books.

Krosnick, J. A., & Brannon, L. A. (1993). The impact of the gulf war on the ingredients of presidential evaluations: Multidimensional effects of political involvement. *American Political Science Review*, *87*, 963–975.

Krosnick, J. A., & Kinder, D. R. (1990). Altering the foundations of support for the president through priming. *American Political Science Review*, *84*, 497–512.

Majone, G. (1989). *Evidence, arguments, and persuasion in the policy process*. New Haven, CT: Yale University Press.

Mark, M. M. (2001). Evaluation's future: Furor, futile, or fertile? *American Journal of Evaluation*, *22*, 457–479.

Mark, M. M. (2003). Toward a comprehensive view of the theory and practice of program and policy evaluation. In S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the new millennium* (pp. 183–204). Hillsdale, NJ: Lawrence Erlbaum.

Mark, M. M., & Henry, G. T. (in press). *The mechanisms and outcomes of evaluation influence*.

Mark, M. M., Henry, G. T., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. San Francisco, CA: Jossey-Bass.

McCombs, M. E., & Zhu, J. (1995). Capacity, diversity, and volatility of the public agenda: Trends from 1954 to 1994. *Public Opinion Quarterly*, *59*, 495–525.

Mintrom, M. (1997). Policy entrepreneurs and the diffusion of innovation. *American Journal of Political Science*, *41*, 738–770.

Monroe, A. (1998). Public opinion and public policy: 1980–1993. *Public Opinion Quarterly*, *62*, 6–27.

Moscovici, S., Lage, E., & Naffrechoux, M. (1969). Influence of a consistent minority on the responses of a majority to a color-perception task. *Sociometry*, *32*, 365–380.

Nelson, T., Clawson, R. A., & Oxley, Z. M. (1997). Media framing of a civil liberties conflict and its effect on tolerance. *American Political Science Review*, *91*, 1002–1009.

Neuman, R. W. (1990). The threshold of public attention. *Public Opinion Quarterly*, *54*, 159–176.

Page, B. I., & Shapiro, R. Y. (1992). *The rational public.* Chicago, IL: The University of Chicago Press.

Patton, M. Q. (1988). The evaluator's responsibility for utilization. *Evaluation Practice*, *9*, 5–24.

Patton, M. Q. (1997). *Utilization-focused evaluation: The new-century text.* Thousand Oaks, CA: Sage.

Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (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* (pp. 141–164). Lexington, MA: Lexington Books.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude.* New York, NY: Springer-Verlag.

Preskill, H., & Caracelli, V. J. (1997). Current and developing conceptions of use: Evaluation Use Topical Interest Group survey results. *Evaluation Practice*, *18*, 209–225.

Reynolds, A. J., Temple, J. A., Roberston, D. L., & Mann, E. A. (2001). Long-term effects of an early childhood intervention on educational achievement and juvenile arrest. *Journal of the American Medical Association*, *285*, 2339–2346.

Rog, D. J. (1985). *A methodological analysis of evaluability assessment.* Doctoral dissertation, Vanderbilt University, Nashville, TN.

Russ-Eft, D., Atwood, R., & Egherman, T. (2002). Use and non-use of evaluation results: Case study of environmental influences in the private sector. *American Journal of Evaluation*, *23*, 33–44.

Sabatier, P. A., & Jenkins-Smith, H. C. (1993). *Policy change and learning: An advocacy coalition approach.* Boulder, CO: Westview Press.

Sabatier, P. A., & Jenkins-Smith, H. C. (1998). The advocacy coalition framework. In P. A. Sabatier (Ed.), *Theories of policy change* (pp. 117–166). Boulder, CO: Westview Press.

Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: Theory, research, and practice since 1986. *Evaluation Practice*, *18*, 195–208.

Shulock, N. (1999). The paradox of policy analysis: If it is not used, why do we produce so much of it? *Journal of Policy Analysis and Management*, *18*, 226–244.

Sinclair, R. C., Mark, M. M., & Shotland, R. L. (1987). Construct accessibility and generalizability across response categories. *Personality and Social Psychology Bulletin*, *13*, 239–252.

Suchman, E. A. (1967). *Evaluation research: Principles and practices in public service and social action programs.* New York, NY: Russell Sage Foundation.

Weick, K. E., & Quinn, R. E. (1999). Organizational change and development. *Annual Review of Psychology*, *50*, 361–386.

Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review*, *39*, 426–431.

Weiss, C. H. (1988a). Evaluation for decisions: Is anybody there? Does anybody care? *Evaluation Practice*, *9*, 5–19.

Weiss, C. H. (1988b). If program decision hinged only on information: a response to Patton. *Evaluation Practice*, *9*, 15–28.

Weiss, C. H. (1998). Improving the use of evaluations: Whose job is it anyway? *Advances in Educational Productivity*, *7*, 263–276.

Weiss, C. H., & Bucuvalas, M. J. (1977). The challenge of social research to decision making. In C. H. Weiss (Ed), *Using social research in public policy making* (pp. 213–234). Lexington, MA: Lexington Books.

Zaslow, M. J., Moore, K. A., Brooks, J. L., Morris, P. A., Tout, K., Redd, Z. A., & Emig, C. A. (2002). Experimental studies of welfare reform and children. *The Future of Children*, *12*, 79–95.

Zhu, J. (1992). Issue competition and attention distraction: A zero-sum theory of agenda setting. *Journalism Quarterly*, *69*, 825–836.