



Evaluation Use

Results From a Survey of U.S. American Evaluation Association Members

Dreolin N. Fleischer

Christina A. Christie

Claremont Graduate University

This paper presents the results of a cross-sectional survey on evaluation use completed by 1,140 U.S. American Evaluation Association members. This study had three foci: evaluators' current attitudes, perceptions, and experiences related to evaluation use theory and practice, how these data are similar to those reported in a previous study conducted by Preskill and Caracelli (1997), and to identify characteristics that distinguish high endorsers of use items from others in the sample. Findings suggest a fair level of agreement on several dimensions of use including stakeholder involvement, factors that influence use, and the varied roles of the evaluator. Logistic regression results indicated that external evaluators were less likely to be high item endorsers while those who reported being members of the Evaluation Use Topical Interest Group of the American Evaluation Association were more likely to be high item endorsers.

Keywords: *evaluation use; stakeholder involvement; organizational outcomes; role of the evaluator; misuse*

A primary objective of evaluation is to provide meaningful information from which decisions about programs and related policies can be made (House, 1980; Patton, 1997a; Stufflebeam & Shinkfield, 1985; Worthen & Sanders, 1973). The intent is that decision makers use the information provided by evaluations to inform their decision making. Earlier scholars recognized that evaluations were not being used, and this observation inspired many to discuss and investigate types of use as well as the factors and conditions that affect use (Alkin, Daillak, & White, 1979; Alkin & Taut, 2003; Patton et al., 1977; Weiss, 1972). Consequently, a body of literature now exists that encompasses varying views about use and calls are periodically made to increase the body of research on evaluation (Alkin, 2003; Shadish, Cook, & Leviton, 1991; Smith, 1993). Our motivation for conducting the current study was to contribute to the empirical research on use and potentially guide future research with our results.

Approximately 10 years ago, Preskill and Caracelli (1997) conducted a study exploring the linkages between theories of evaluation use and perspectives on practice. The field of evaluation has shifted and grown since Preskill and Caracelli's study, yet the topic of evaluation use

Authors' Note: Dreolin N. Fleischer, Claremont Graduate University, 123 E. 8th St., Claremont, CA 91711; e-mail: dreolin.fleischer@cgu.edu. We wish to thank Michael Patton for suggesting that we conduct this particular study; Hallie Preskill and Valerie Caracelli for so readily allowing us to adapt their earlier survey; Hallie Preskill for her continued support throughout the research and writing process; and the many individuals who gave feedback on the instrument revision. We would also like to thank AEA for allowing us to do the study with their members and Robin Miller and our anonymous reviewers for their extremely valuable feedback on this manuscript.

has remained relevant and of interest to academics and practitioners alike. Our study offers a current description of evaluators' views on the issue of evaluation use. Preskill and Caracelli's study provides a historical point of reference from which to consider the current data. The current study addressed the following three foci: (a) The current attitudes, perceptions, and experiences of U.S. members of the American Evaluation Association (AEA) on the topic of evaluation use in theory and practice; (b) How those attitudes, perceptions, and experiences are similar to or different from those reported in the Preskill and Caracelli's study; and (c) Respondent characteristics, specifically in reference to their endorsement of clusters of items revealed through factor analysis. In addressing the three study foci, there were five domains of particular interest: factors that influence use, evaluation as it relates to organizational outcomes, the role of the evaluator, stakeholder involvement, and misuse. The current study, like the 1996 study, was exploratory and did not have a priori hypotheses.

Topics Related to Evaluation Use: Support for the Study Domains

It is common for evaluators to anticipate that their work will be useful to policy makers, program managers, and other stakeholders in solving social problems (Shadish, Cook, & Leviton, 1991). Shadish et al. (1991) argue that society justifies spending large amounts of money for evaluations with the expectation that there will be some immediate payoffs; thus, if evaluations are not useful, those funds should be expended for alternative purposes. Consequently, evaluation use has been a topic of concern for evaluation scholars for some time. Unlike many other evaluation topics, systematic study of evaluation use has led to the development of empirically grounded conceptualizations of use. This section defines various types of evaluation use and provides background on each of the study domains addressed in the 2006 survey.

Definitions of Use

Over the past several decades, evaluation scholars have created terminology and definitions for different types of evaluation use (Kirkhart, 2000; Patton, 1994; Weiss, 1980). Five types of use are commonly found in the literature. The first three are primarily related to the use of findings: instrumental, when decision makers use the evaluation findings to modify the evaluand (i.e., the object of the evaluation) in some way (Greene, 1988; Rich, 1977; Shulha & Cousins, 1997); conceptual, when the evaluation findings help program staff understand the program in a new way (Rich, 1977; Weiss, 1979); and enlightenment, when the evaluation findings add knowledge to the field and thus may be used by anyone, not just those involved with the program or evaluation of the program (Weiss, 1979).

In addition to the three types of use specified above, Patton (1997a, 2003) describes a fourth type of use—process use. Process use is defined as “. . . cognitive, behavioral, program and organizational changes resulting from engagement in the evaluation process and learning to think evaluatively” (Patton, 2003, p. 230). The fifth type of use is persuasive or symbolic use (Alkin & Taut, 2003; Greene, 1988; Henry, 2000; Leviton & Hughes, 1981; Weiss, 1998). More so than the other definitions of use, there is a lack of consensus on what should be considered persuasive or symbolic use. Some consider using an evaluation to persuade important stakeholders that the program or organization values accountability as an example of persuasive or symbolic use. A less desirable example of persuasive or symbolic use is when an evaluator is hired to evaluate a program to legitimize a decision that has already been made prior to the commissioning of the evaluation.

Planning for Use: Factors That Facilitate Use

Much of the extant research on evaluation use focuses on how and under what conditions evaluation findings are used. Cousins and Leithwood's (1986) article on the use of evaluation presents the results from 65 empirical studies conducted between 1971 and 1986. Using a meta-analytic method to determine which factors more heavily affect the use of evaluation data, Cousins and Leithwood found that the studies showed evaluation use is encouraged by a user-focused evaluation approach. User-focused evaluation approaches, where use was strongly evident, shared several features. To begin, the approach, methodological sophistication, and intensity of the evaluation were perceived to be appropriate by program stakeholders. Users were also involved in the evaluation process and invested in the benefits of evaluation. Another characteristic was that these evaluations produced relevant findings that were consistent with the previously held beliefs and expectations of the stakeholders. Finally, other sources rarely conflicted with the evaluation findings (Cousins & Leithwood, 1986, p. 360).

More recently, Preskill, Zuckerman, and Matthews (2003) conducted an exploratory study of process use. The study identified several factors that appear to affect process use. The five categories of factors were as follows: (a) facilitation of evaluation processes; (b) management support; (c) advisory group characteristics; (d) frequency, methods, and quality of communications; and (e) organization characteristics. The authors described the concept of process use as being "...based on social constructivist learning theory which suggests that individuals construct knowledge and develop a shared reality through collaboration with others" (p. 424). An empirical research base on this topic is growing. Amo and Cousins (2007) reviewed 18 studies related to process use published between 1984 and 2005. Theorized linkages between process use, evaluation capacity building, and organizational learning were revealed by the review and suggest process use is still quite relevant within the field of evaluation.

Evaluation as It Relates to Organizational Outcomes

The study by Preskill and Caracelli (1997) reflected a changing interest in the body of use literature at that time. Theorists and researchers began to look beyond factors that facilitate instrumental use of evaluation findings to the issue of process use (Patton, 1997a) and that led authors to examine, among other things, how process use can be a mechanism for transformative learning within organizations. Preskill and Torres (1999, 2000) led efforts to strengthen the connection between evaluation and organizational learning through a collaborative approach to evaluation, grounded in constructivist and transformative learning theories.

As the relationship between evaluation and organizational outcomes becomes more recognized, efforts to study the topic systematically have begun. In 2004, Cousins, Goh, Clark, and Lee reviewed the literature relevant to evaluation utilization, evaluation capacity building, and organizational learning. The authors viewed all three concepts as being interrelated. From 36 recent empirical studies, they developed a conceptual framework that depicted how evaluation capacity overlaps with organizational learning capacity to create what is referred to as "organizational readiness for evaluation." An expanded conceptual model is currently in development (Cousins, Goh, & Elliot, 2007); it proposes a set of complex linkages among multiple concepts related to evaluation capacity building (e.g., organizational support structures, capacity to *do* evaluation, evaluative inquiry, etc.).

The Role of the Evaluator in Facilitating Use

Several factors can influence the role of an evaluator, including the purpose of the evaluation, stakeholders' information needs, the evaluator's epistemological preferences, and the evaluation

approach used. Preskill and Caracelli (1997) found that 95% of respondents thought that “evaluators should take responsibility for involving stakeholders in the evaluation process” (p. 221). The increased interest in participatory and collaborative evaluation approaches has shifted the role of evaluator away from the traditional notion of a detached evaluator (i.e., an evaluator who limits their interaction with stakeholders in an effort to maintain an unbiased point of view). In participatory evaluation, the evaluator may be expected to be an “evaluator, facilitator, critical friend, organization developer, mediator, and the like” (Cousins, 2003, p. 261). Rossman and Rallis (2000) also maintain that within a learning-based evaluation approach, the evaluator’s role is of “critical friend” or “partner and co-producer of knowledge” (p. 67).

Stakeholder Involvement

Stakeholder involvement is an important aspect of participatory and collaborative evaluation approaches. However, attention to the involvement of stakeholders extends beyond the aforementioned approaches; indeed, it is accepted more generally that stakeholder involvement increases evaluation use (Christie, 2003; Greene, 1988; Turnbull, 1999). Reineke (1991) suggests that quality involvement is a result of the evaluator identifying the stakeholders early on and then continuously engaging them, while establishing a structure for dialogue. Stakeholder involvement has become a mainstream concept in evaluation as is evident in its role within most evaluation approaches. However, there is a relative dearth of empirical research on this topic, as compared with the theoretical or applied writings on the subject.

Misuse of Evaluation Findings

The issue of misuse has received relatively little attention in the literature. Alkin and Coyle (1988) and then later Christie and Alkin (1999) attempt to describe evaluation misuse and the various dimensions related to it. The dimensions include intentional (e.g., using the evaluation process to stall decision making) and unintentional misuse (e.g., when the client is genuinely unaware that their actions are leading to misuse of the evaluation findings) as well as nonuse. Evaluators cannot always prevent misuse of evaluation findings but should be held responsible for misevaluation. “In essence, misevaluation is designing an evaluation with flawed methodology, careless data collection, sloppy analysis, and/or poor reporting” (Christie & Alkin, 1999, p. 6). Moreover, nonuse of the evaluation is considered appropriate when misevaluation has occurred. If the client is aware that misevaluation has occurred, but persists in using the evaluation findings, this is considered misuse. As demonstrated in the literature review above, evaluation use remains a topic of interest for scholars in the field. These five domains represent core topics within the literature on evaluation use and thus provided the organizational structure of our study. In 2006, we conducted a survey of U.S. members of AEA, and it is our belief that the results of this survey contribute some new and meaningful information about these topics to the current field of evaluation. The remainder of this article describes the study methods and results.

Method

The three study foci and five study domains were addressed using a cross-sectional survey design. The primary strength of this design is that it allows researchers to collect data from a large number of people at one time. Another reason this design was used is because it matched

the design Preskill and Caracelli (1997) used for their study, allowing us to make comparisons more easily between the two studies. We realize that comparing results from two different surveys, on two different samples, at two different time points is not ideal; however, it is useful in being able to piece together a sense of where the field was 10 years ago and where it is today. Understanding these changes, or in some cases lack of changes, over time can help evaluators and researchers reflect on the state of the field. Time spent reflecting, as part of a cycle of research and dialogue, is key to developing the field's identity as well as promoting future avenues of research on the topic of evaluation use.

Procedure

In 2006, after undergoing an application process with AEA's Executive Board, we obtained the names and e-mail addresses of 3,919 U.S. AEA members. Of them, 95 members had no e-mail address, an incomplete e-mail address, or a duplicate e-mail address of some other member. As a result, an invitation to participate with a link to the online survey was successfully e-mailed to 3,824 members on May 25, 2006; of those, 1,140 members completed a survey for a response rate of 30%. Just over half of the respondents completed the survey prior to the distribution of the reminder e-mail on June 1, 2006. The 2006 data collection procedure differed from that of the 1996 study. Preskill and Caracelli (1997) successfully mailed their paper and pencil survey to 525 North American AEA members who, at that time, were also members of the Evaluation Use Topical Interest Group (TIG). Ultimately, 282 Use TIG respondents completed and returned the survey for a response rate of 54%.

Participants

In 2006, the majority of respondents (96%) reported that evaluation was part of their professional identity in some form (i.e., either their primary or secondary professional identity). At the time of the survey, 92% of respondents were currently conducting evaluations; 37% identified as an external evaluator, 22% identified as an internal evaluator, and 33% identified as both an external and internal evaluator. The majority of respondents (87%) classified themselves as either having an intermediate or advanced level of knowledge of, and experience in, evaluation. Respondents reported differing levels of experience with a majority of respondents (57%) reporting that they had spent 10 or fewer years conducting evaluations. Half of the respondents reported working on a combination of both small and large-scale evaluations. For those who typically worked on either small or large-scale evaluations, more participants worked on small-scale evaluations (34%) than large-scale evaluations (14%). Of those respondents conducting evaluations, on average, 39% of respondents conduct 1–3 evaluations per year, 30% conduct 4–6 evaluations per year, and 26% conduct 7 or more per year.

In all, 22% of participants reported that they were members of AEA's Evaluation Use TIG, 55% reported not being a member of the Evaluation Use TIG, and 24% reported "I don't know/don't remember" to the question. The 2006 sample of U.S. AEA members closely resembled the 1996 sample of AEA Evaluation Use TIG members on all background questions despite the differences in sample populations: only Evaluation Use TIG members were surveyed in 1996, whereas the larger U.S. membership of AEA was surveyed in 2006.

Measures

As stated previously, the 1996 and 2006 surveys were not identical. The 2006 survey needed to be similar enough to facilitate comparisons between it and the 1996 survey but

different enough to allow for new information to be collected and for some language and concepts to be updated to better reflect the current field of evaluation. We also shortened the survey to 73 items and reversed all Likert-type response scales to read from left to right “strongly disagree” to “strongly agree,” so that the negative anchor was positioned on the left and the positive anchor was positioned on the right (mimicking a number line) to make the scale familiar to respondents, thus reducing the cognitive demand necessary for completing the survey. Additionally, we reworded the scale for the factors that influence use items so that respondents rated the influence (i.e., “not at all influential” to “extremely influential”) rather than the importance of each factor.

We collected background data on respondents’ knowledge about evaluation, evaluator role (internal or external evaluator), type of organization in which they predominantly work, the number of evaluations conducted per year, the size of evaluations typically conduct (small or large scale), the length of time they have been conducting evaluations, and Evaluation Use TIG membership. Respondents were then asked the extent to which they agree or disagree with 10 statements on various roles evaluators might play in an evaluation. Sample items were “maximizing intended use by intended users” and “facilitating organizational learning.”

Consistent with the 1996 survey, a definition for use of evaluation findings was offered to provide context for the remainder of survey items which all focused on evaluation use. Seven statements focused on the nature and extent of stakeholder involvement in an evaluation. These items included “involving multiple stakeholders increases the use of evaluation findings” and “involving multiple stakeholders increases process use.” Respondents were asked to rate the influence of 15 activities that might have an impact on evaluation use. Items included “developing a communicating and reporting plan” and “planning for use at the beginning of the evaluation.” Ten statements focused on the extent to which organizational outcomes can be fostered by evaluation. Examples of these items include “plays a pivotal role in transforming organizations” and “develops users’ systematic inquiry skills.” Respondents were asked to report the extent to which they believe “nonuse of evaluation results,” “intentional misuse of evaluation results,” and “unintentional misuse of evaluation results” were a problem and then asked to rate the frequency with which they see six examples of misuse, including seeing clients “dismiss undesirable results,” “selectively report results,” and “exaggerate results.”

Results

Descriptive Results

In this section, the results from the entire 2006 sample will be presented first, followed by a comparison between the 1996 Evaluation Use Topical Interest Group (Use TIG) responses and the 2006 Use TIG responses. We will conclude the section by exploring differences between the 2006 Use TIG responses and those of the 2006 non-Use TIG responses. We chose to report the percentage of respondents who positively endorsed the items, so that the 2006 results could be more easily compared with the 1996 results. The results of the 1996 study were mainly reported as percentages within Preskill and Caracelli’s (1997) article.

Results From the 2006 Sample

Factors that influence evaluation use. Respondents were asked to rate, on a scale of 1 “not at all influential” to 5 “extremely influential,” the influence of 15 factors in increasing

Table 1
Respondents' Perspectives Concerning Factors That Influence Use

	Influential to Extremely Influential (%), <i>n</i> = 897
Planning for use at the beginning of the evaluation	91
Identifying and prioritizing intended <i>uses</i> of the evaluation	89
Communicating findings to stakeholders as the evaluation progresses	87
Identifying and prioritizing intended <i>users</i> of the evaluation	86
Involving stakeholders in the evaluation process	86
Developing a communicating and reporting plan	86
Interweaving the evaluation into organizational processes and procedures	83
Soliciting and using suggestions from stakeholders on the evaluation design	76
Designing the evaluation within resource limitations	74
Linking the evaluation to an established pattern of organizational life	72
Adhering to high standards of methodological rigor	65
Maintaining significant involvement in evaluation follow-up activities	65
Taking steps to prevent the misuse of evaluation findings	61
Demonstrating that the benefits of the evaluation will outweigh the costs	57
Establishing a balance of power among stakeholders	44

Note: The sample size varies across items. We chose to present the smallest sample size from the series of items as a conservative representation. This is the case for every table in the article.

evaluation use (see Table 1). These factors are not mutually exclusive; an evaluator could attempt to follow some or even all of these in any given evaluation. Depending on the factor, the evaluator, and the type of evaluation being conducted, varying amounts of time and effort would be needed to accomplish each. In addition, addressing each factor adequately would require a competent evaluator with an extensive and varied skill set.

Perhaps not surprisingly, factors having to do with use were seen as most influential in ultimately increasing the likelihood of use occurring. Of particular interest in Table 1 is the item related to methodological rigor. Research on use suggests that for use to occur the evaluation must be credible (e.g., Patton, 1997a). Methodological rigor often is portrayed as a key component of to being able to offer credible evidence (e.g., Bickman & Reich, 2009; Gersten & Hitchcock, 2009; Henry, 2009; Julnes & Rog, 2009). Yet, adherence to methodological rigor was located in the bottom third of the items when organized from highest to lowest according to the level of endorsement.

Organizational outcomes. Respondents were asked to rate, on a scale from 1 “no extent” to 5 “great extent,” the extent to which they believe evaluation fosters 10 separate organizational outcomes (see Table 2).

Respondents rated “enhances group learning” slightly higher than “enhances individual learning” when considering the extent to which evaluation fosters either of these organizational outcomes. In a noticeable departure from the higher ratings of the other nine items, relatively few respondents reported that evaluation “changes social power structures.” The practical significance of this finding varies according to the extent to which one personally believes changing social power structures is part of the scope of evaluation.

The role of the evaluator. Respondents were asked to rate, on a scale from 1 “strongly disagree” to 5 “strongly agree,” the extent to which they agreed evaluators should take responsibility for 10 separate evaluation activities (see Table 3).

Table 2
Respondents' Perspectives Concerning the Role of Evaluation in Fostering Organizational Outcomes

	Some Extent to a Great Extent (%), n = 912
Contributes to an organization's ability to learn from its experiences	84
Can be a powerful planned change strategy	78
Helps intended users question basic assumptions about their practice	73
Increases users' evaluative thinking	71
Enhances group learning	69
Enhances individual learning	66
Develops users' systematic inquiry skills	52
Plays a pivotal role in transforming organizations	45
Results in fundamental changes in how organizations achieve their goals	38
Changes social power structures	14

Table 3
Respondents' Perspectives Concerning the Role of the Evaluator

	Agree to Strongly Agree (%) n = 1,047
Involving stakeholders in the evaluation process	98
Facilitating organizational learning	89
Maximizing intended use by intended users	86
Formulating recommendations from evaluation studies	86
Building evaluation capacity	82
Being accountable to intended users of the evaluation for intended uses of the evaluation	77
Becoming involved in evaluation follow-up activities	64
Attending to the social consequences of the evaluation	64
Making the interests of the economically and politically disadvantaged a priority	56
Becoming program advocates	17

Although the reported level of agreement was fairly high for most of the activities, involving stakeholders in the process was the most agreed on role of the evaluator. Interestingly, facilitating organizational learning was the second most agreed on role of the evaluator. This result points to the growing awareness within the field of the connection between evaluation and organizational learning. The role of the evaluator that was endorsed by the fewest respondents was becoming program advocates, suggesting that program advocacy is largely not considered a responsibility of an evaluator.

Stakeholder involvement. Respondents were asked to rate, on a scale from 1 "strongly disagree" to 5 "strongly agree," the extent to which they agreed with seven statements about stakeholder involvement in the evaluation (see Table 4). A larger percentage of respondents agreed "stakeholder participation is most important when focusing the evaluation" than agreed "it is important that stakeholders participate in all stages of the evaluation." These data suggest stakeholder involvement at the start of the evaluation is more widely considered important than involving stakeholders at all stages of the evaluation.

Table 4
Respondents' Perspectives Concerning Stakeholder Involvement

	Agree to Strongly Agree (%) <i>n</i> = 954
Some key stakeholders reject conclusions on the basis of their own "beliefs and values" rather than on the data	86
Involving multiple stakeholders increases the use of evaluation findings	78
Stakeholder participation is most important when focusing the evaluation	74
Involving multiple stakeholders increases process use	70
Involving multiple stakeholders helps balance political agendas	65
It is important that stakeholders participate in all stages of the evaluation	48
Stakeholders as decision-makers rarely base their decisions directly on evaluation information	46

It is also interesting to note that although the 2006 respondents agreed overwhelmingly that stakeholders have a key role in the evaluation process, items indicating that stakeholders do not use evaluation findings appropriately—if at all—were also highly endorsed. Respondents endorsed the statement "involving multiple stakeholders increases the use of evaluation findings," yet 86% agreed "some key stakeholders reject conclusions on the basis of their own 'beliefs and values' rather than on the data." Far fewer, but still a noteworthy group, also agreed that "stakeholders as decision makers rarely base their decisions directly on evaluation information."

Misuse and nonuse. Respondents were asked to report on the extent to which nonuse, intentional misuse, and unintentional misuse of evaluation results were a problem based on their personal experiences. Nonuse of evaluation results was considered "a major problem" by 68% of respondents, whereas intentional misuse of evaluation results and unintentional misuse of evaluation results were seen as "a minor problem" by 52% and 66%, respectively.

Use TIG Members Comparison 1996–2006

Although the overall composition of the AEA membership has undoubtedly changed over the 10-year period, it is likely that at least some of the individuals who completed the 1996 survey also completed the 2006 survey. Comparisons between the results of the two surveys are meant only to highlight differences and similarities in cross-sectional samples of Use TIG members' responses at different points in time. In this section, we only highlight the notable differences in responses between the two samples.

Factors that influence evaluation use (refer to Table 1 for survey items). For the majority of the items, a slightly greater percentage of 2006 respondents positively endorsed factors that influence evaluation use; specifically, in 1996, 77–90% of the sample positively endorsed seven items compared with 82–94% in 2006 (for the same items). Slightly fewer 2006 respondents positively endorsed one item ("designing the evaluation within resource limitations"). There were two items ("demonstrating the benefits of evaluation will outweigh the costs" and "establishing a balance of power among stakeholders") on which 38% or fewer of the 1996 sample endorsed positively. Although these same two items were also positively endorsed by the fewest number of respondents on the 2006 survey, a larger proportion of the 2006 sample rated these items positively (55% or fewer).

Organizational outcomes (refer to Table 2 for survey items). Eight of the 10 organizational outcome items were identical to those asked in the 1996 survey. In 2006, a greater proportion of Use TIG respondents (ranging from 14% to 25% more in 2006, with an average increase of 21%) considered the eight organizational outcomes as being fostered by evaluation. It seems likely that expanding roles for evaluators will include facilitator of organizational learning and change.

The role of the evaluator (refer to Table 3 for survey items). Few respondents in both the 1996 and 2006 samples agreed that it was the role of the evaluator to be a program advocate (15% in 1996 and 20% in 2006). A slightly greater proportion of respondents agreed that “making the interests of the economically and politically disadvantaged a priority” is an evaluator’s responsibility, although a larger percentage of the 2006 sample (59%) endorsed this item than the 1996 sample (39%).

Stakeholder involvement (refer to Table 4 for survey items). When looking at the 2006 and 1996 samples, overall there was little difference in opinion across the two time points. Of particular interest was the consistent view by the great majority of the respondents in both 1996 (89%) and 2006 (85%) that “some key stakeholders reject conclusions on the basis of their own ‘beliefs and values’ rather than on the data.”

Misuse and nonuse. We found that a greater percentage (74%) of Use TIG members in 2006 reported “nonuse of evaluation results” to be “a major problem” (46% in 1996). A greater percentage (29% and 33%, respectively) of Use TIG members in 2006 reported “intentional misuse of evaluation results” and “unintentional misuse of evaluation results” to be “a major problem” (26% and 24%, respectively, in 1996).

2006 Use TIG and Non-Use TIG Comparison

Differences were expected between Use TIG members and non-Use TIG members because membership in the Evaluation Use TIG was assumed to suggest some level of interest and possible investment in the topic of the TIG.

Factors that influence evaluation use (refer to Table 1 for survey items). On all but two items, 2006 Use TIG members rated the factors more favorably than 2006 non-Use TIG members; this difference was statistically significant (at either $p < .05$ or $p < .01$) for five of the items: “planning for use at the beginning of the evaluation,” “identifying and prioritizing intended users of the evaluation,” “identifying and prioritizing intended uses of the evaluation,” “communicating findings to stakeholders as the evaluation progresses,” and “establishing a balance of power among stakeholders.” These results suggest respondents who identified as more use-focused had a more optimistic impression of the extent to which these factors influenced use.

Organizational outcomes (refer to Table 2 for survey items). On all 10 items, 2006 Use TIG members rated the items more favorably than 2006 non-Use TIG members; this difference was statistically significant (at either $p < .05$ or $p < .01$) for five of the items: “contributes to an organization’s ability to learn from its experiences,” “helps intended users question basic assumptions about their practice,” “enhances group learning,” “increases users’ evaluative thinking,” and “plays a pivotal role in transforming organizations.” These results suggest that

Use TIG members were more likely to think these organizational outcomes could be fostered by evaluation than non-Use TIG members.

The role of the evaluator (refer to Table 3 for survey items). On 9 of the 10 items, 2006 Use TIG members rated the items more favorably than 2006 non-Use TIG members; this difference was statistically significant (at either $p < .05$ or $p < .01$) for five of the items: “involving stakeholders in the evaluation process,” “facilitating organizational learning,” “maximizing intended use by intended users,” “building evaluation capacity,” and “attending to the social consequences of the evaluation.” The pattern of responses suggests that Use TIG members were more likely to see their role as evaluator more multidimensional than non-Use TIG members.

Stakeholder involvement (refer to Table 4 for survey items). On five of the seven items, 2006 Use TIG members rated the items more favorably than 2006 non-Use TIG members; this difference was statistically significant (at either $p < .05$ or $p < .01$) for two of the items: “it is important that stakeholders participate in all stages of the evaluation” and “stakeholders as decision-makers rarely base their decisions directly on evaluation information.”

Misuse and nonuse. In 2006, non-Use TIG members were significantly less likely to see nonuse as a problem when compared with Use TIG members at a statistically significant level ($p < .05$). Non-Use TIG members, in 2006, were less likely to consider intentional misuse of evaluation results and unintentional misuse of evaluation results as a problem at a statistically significant level ($p < .01$).

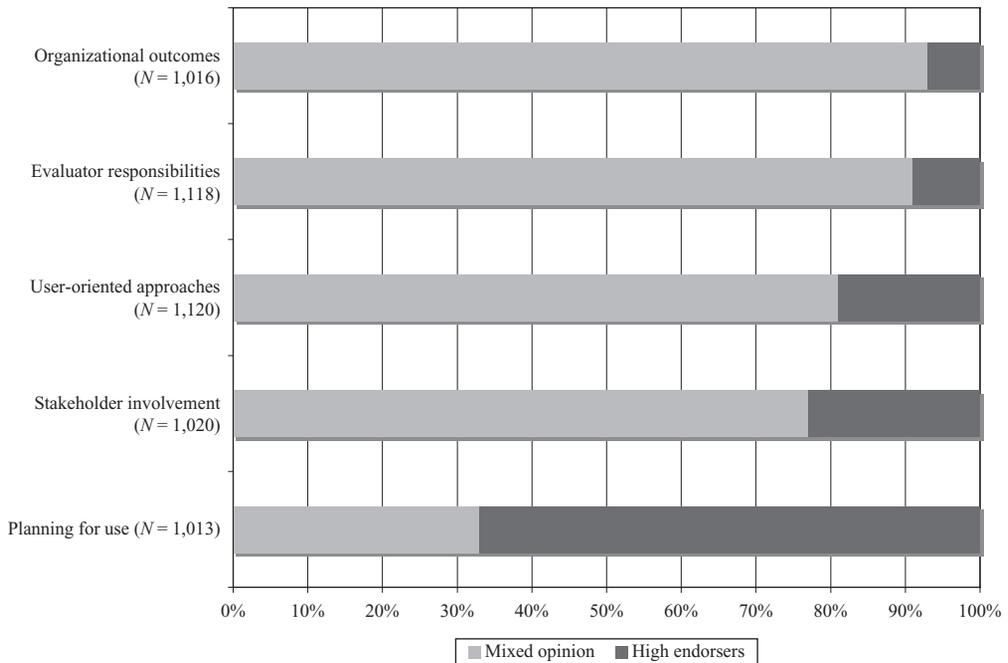
Analysis of the “High Endorsers”

As a compliment to the descriptive results, additional statistical analyses were conducted. First, a factor analysis was carried out to reveal possible underlying constructs among the survey items. Essentially, the goal was to reduce the data in a meaningful way for the purpose of including only those items that clustered together as factors in further statistical analyses. Second, in an effort to explore potential differences between respondents, each respondent was categorized as either a “high endorser” (i.e., those who had strongly held positive beliefs about aspects of evaluation use) or of a “mixed opinion” based on the factor analytic results. The criteria used to categorize respondents into these two groups will be described below. Third, a series of binary logistic regression analyses were conducted to determine whether certain evaluator characteristics predicted membership into the “high endorser” group for any of the factors revealed through the factor analysis.

Factor Analysis

Responses to all of the Likert-type scale items on the 2006 survey were factor analyzed using a principal components analysis (PCA) with a varimax rotation that resulted in a five-factor solution retaining 29 of the 50 items. The optimal solution was decided based on the Kaiser rule, amount of variance explained, and the fact that these five factors made theoretical sense. Factor 1 included 10 items ($\alpha = .880$) and was given the name “organizational outcomes” because it included each of the 10 items grouped together on the survey about the ability of evaluation activities to foster organizational outcomes. Factor 2 included six items ($\alpha = .757$) and was given the name “stakeholder involvement” because it consisted of items from three separate portions of the survey that asked about the manner in which stakeholders

Figure 1
Mixed Opinion and High Endorsers According to Factor



should be included in the evaluation. Factor 3 included four items ($\alpha = .733$) and was given the name “planning for use” because it consisted of items from the section of the survey that asked respondents to rate the extent to which certain factors can increase evaluation use. Factor 4 included five items ($\alpha = .807$) and was given the name “user-oriented approaches” because it included the five of the user-oriented approach items from the series of items that asked respondents to rate the importance of a variety of different evaluation approaches. Finally, Factor 5 included four items ($\alpha = .638$) and was given the name “evaluator responsibilities” because it consisted of items from the section of the survey that asked respondents to indicate whether they agree or disagree with various roles the evaluator can take within an evaluation.

High Endorsers

A “high endorser” gave every item included in a factor either a “4” or a “5” on the five-point Likert-type response scale, whereas a person of “mixed opinion” gave a more varied response to the items within each factor. To give a sense of the prevalence of “high endorsers,” the overall proportion of “mixed opinion” respondents to “high endorser” respondents for each factor can be seen in Figure 1.

Logistic Regression Analysis

A series of logistic regression models were run to see whether four dummy coded background items predict membership into the high endorser group for each of the five factors.

Descriptive information about the evaluator, including the role of the evaluator (*External*: 1 = external; 0 = internal), the size of the evaluations they typically work on (*Large scale*: 1 = large scale; 0 = small scale), years of experience (*11 or more years*: 1 = 11 or more; 0 = 10 or less), and whether they were members of the Use TIG (*TIG Member*: 1 = members; 0 = not members), were used as predictor variables (variable names are italicized). Outcome variables included each of the five factors yielded from the factor analysis.

It is important to note that for three of the predictor variables mentioned above (*External*, *Large scale*, and *TIG Member*) the variables were coded in such a way that portions of the respondents were not included in the analyses. In the case of *External*, only respondents who reported that they primarily conduct evaluations as an external evaluator (37% of the sample) and respondents who reported that they primarily conduct evaluations as an internal evaluator (22% of the sample) were included in the logistic regression analyses. Respondents who reported that they act as both internal and external depending on the evaluation (33% of the sample) and respondents who reported that they were not currently conducting evaluations (8% of the sample) were not included. In the case of *Large scale*, only respondents who reported typically working on large-scale evaluations (14% of the sample) and respondents who reported typically working on small-scale evaluations (34% of the sample) were included and respondents that said they work on a combination of both (50% of the sample) or reported "other" (3% of the sample) were not included. For TIG member, those individuals who reported "I don't know/Don't remember" (24% of the sample) were not included in the logistic regression (although they are included as part of non-Use TIG members for all other analyses). Although the reduction in observations is not ideal, the restriction allowed for more meaningful and interpretable results. Moreover, the remaining sample size was still sufficient to conduct the analyses (over 200 cases).

A forced-entry logistic regression rather than a hierarchical logistic regression was run for each of the five models (one per factor) because the analyses were exploratory. Logistic regression does not require normal distributions, so none of the variables needed to be transformed. The assumptions of logistic regression analyses are that the data are from a random, independent sample and that there is linearity between x and $\log[P/(1-P)]$. There is no previous research that suggests any of the predictor variables are more or less important in predicting membership into the high endorser group for any of these factors; therefore, all the predictors were entered into the model at one time.

Logistic regression results. The regression model was not statistically significant for the evaluator responsibilities outcome variable and therefore is not reported here. However, statistically significant results were found for organizational outcomes, stakeholder involvement, planning for use, and user-oriented approaches outcome variables. First, the two models that yielded two significant predictors will be reported. Second, the two models that yielded one significant predictor each will be reported.

The model predicting membership in the high endorser group for the organizational outcomes factor (overall model chi-square [4, $N = 235$] = 12.87, $p < .05$) yielded two significant predictors. Use TIG members ($B = 1.76$, $SE = 0.77$, Wald $\chi^2 = 5.16$, $p < .05$) are more likely to be high endorsers of the organizational outcomes factor than non-Use TIG members (odds ratio = 5.80; confidence interval [CI] = 1.27–26.44). External evaluators ($B = -1.96$, $SE = 0.85$, Wald $\chi^2 = 5.32$, $p < .05$) are less likely to be a high endorser of the organizational outcomes factor than internal evaluators (odds ratio = 0.14; CI = 0.03–0.75). The model predicting membership in the high endorser group for the planning for use factor (overall model chi-square [4, $N = 234$] = 18.08, $p < .001$) also yielded two significant predictors. Use TIG members ($B = 1.16$, $SE = 0.36$, Wald $\chi^2 = 10.33$, $p < .001$) are more likely to be high endorsers of the planning for use factor than non-Use

TIG members (odds ratio = 3.18; CI = 1.57–6.43). External evaluators ($B = -0.82$, $SE = 0.32$, Wald $\chi^2 = 6.76$, $p < .01$) are less likely to be a high endorser of the planning for use factor than internal evaluators (odds ratio = 0.44; CI = 0.24–0.82).

The model predicting membership in the high endorser group for the stakeholder involvement factor (overall model chi-square [4, $N = 236$] = 6.05, NS) yielded one significant predictor. Use TIG members ($B = 0.74$, $SE = 0.35$, Wald $\chi^2 = 4.61$, $p < .05$) are more likely to be high endorsers of the stakeholder involvement factor than non-Use TIG members (odds ratio = 2.10; CI = 1.07–4.15). The model predicting membership in the high endorser group for the user-oriented approaches factor (overall model chi-square [4, $N = 235$] = 5.19, NS) also yielded one significant predictor. Use TIG members ($B = 0.75$, $SE = 0.36$, Wald $\chi^2 = 4.34$, $p < .05$) are more likely to be high endorsers of the user-oriented approaches factor than non-Use TIG members (odds ratio = 2.12; CI = 1.05–4.29).

Discussion

Modern program evaluation began to take shape in the early 1960s when federal funding for education (e.g., the Elementary and Secondary Education Act of 1965) and antipoverty programs increased, resulting in a demand for evaluations of these programs (Fitzpatrick, Sanders, & Worthen, 2003). Soon thereafter, decision makers began to scrutinize the evaluation studies being conducted, leading to criticism that evaluation findings were not being used (Preskill & Torres, 2000). This criticism spurred research on the topic of evaluation use (Alkin, Daillak, & White, 1979; Alkin & Taut, 2003; Patton et al., 1977; Weiss, 1972). Evaluation use is arguably now the most studied area of evaluation. Yet, the need remains for researchers to revisit and update our notions on the topic of evaluation use. After all, we might expect to find some changes in evaluators' perceptions on use from one decade to the next. Our study provides insight into the current state of evaluation use, examining a sample of evaluators not easily accessed. The AEA only occasionally and cautiously grants researchers permission to contact their membership for research purposes, and our response rate from this group exceeded our expectations.

Given the complexities surrounding evaluation practice and epistemological beliefs, one may expect considerable difference in opinion from a sample of this size. Yet, we found a fair amount of agreement on survey items. Respondents generally agreed that stakeholders should be involved in the evaluation process (at least in some capacity), that there are a number of roles for the evaluator, and that certain evaluation activities can foster organizational learning. There was also a fair level of agreement on the factors that are influential in increasing evaluation use. Looking across sections of the survey, we also identified some noteworthy discrepancies. For example, evaluators endorsed a series of items describing actions an evaluator can take prior to, during, or after an evaluation that can potentially increase use. Yet, elsewhere on the survey 68% of the overall sample and 74% of Use TIG members reported nonuse of evaluation results to be a major problem.

A look into the research on barriers to use (Taut & Alkin, 2003) might help explain this incongruent set of results. There are a variety of factors that can serve as barriers to evaluation use: human factors (relating to user or evaluator characteristics; e.g., knowledge, experience, perceptions), contextual factors (relating to the context surrounding the evaluation; e.g., political background, organizational background), and evaluation factors (relating to how the evaluation is conducted; e.g., ethics, design, methods) (Alkin, 1985). Our data primarily focus on evaluation factors rather than contextual and human factors. Presuming evaluators' practices mirror their endorsement of evaluation factors as influential in increasing use, it is likely

that human and contextual factors are contributing to nonuse more so than evaluation factors. Respondents from the current study did identify two specific barriers to use that lend support to this point. The great majority of the sample agreed that key stakeholders reject conclusions on the basis of their own 'beliefs and values' (human factor) rather than on the data, whereas slightly less than half agreed that decision makers rarely based their decisions directly on evaluation information (context factor).

Although we found evaluators believe that stakeholders might reject or ignore evaluation information, respondents still maintained that stakeholders should be involved in the evaluation process. Because the number of stakeholders as well as the type of stakeholder participation is often determined by the evaluation approach used, this finding supports arguments for including stakeholders in the evaluation process not just for increasing the likelihood of use but also for additional purposes such as helping the evaluator make decisions about what and how to evaluate (Henry & Rog, 1998).

The roles the evaluator assumes during a study is also influenced by the approach used. Some approaches describe the evaluator's role as a "critical friend," whereas others argue that the evaluator should be a critical judge or expert. Role matters because an evaluator's responsibilities are influenced by the roles assumed. Our study findings suggest that evaluators see their roles as wide and varied. A majority of respondents agreed that evaluators were responsible for involving stakeholders in the evaluation process, facilitating organizational learning, maximizing intended use by intended users, and building evaluation capacity, among other things. Our data, however, provide only information about respondents' perceptions of appropriate roles for the evaluator rather than a measure of the roles they actually assume during practice. Complementary information about the roles evaluators actually assume in practice was provided through a study of AEA membership conducted in 2007–2008 (American Evaluation Association, 2008). Using a sample very similar to the one used in the current study (with overlap likely), AEA survey respondents were presented a list of eight evaluation-related work activities and asked to identify all the activities they actually implement in practice. More than half of the respondents reported engaging in technical assistance and evaluation capacity building, whereas slightly less than half of the respondents reported engaging in training others in evaluation (p. 14). Taken together, data from these two surveys offer evidence that the field is continuing to move in a direction where evaluators view themselves as more than the methodological expert.

The one item from the list of 10, related to the evaluator's role, presented on our survey that was not endorsed by a majority of the sample was the role of evaluator as program advocate. A majority of respondents disagreed that becoming program advocates is a role of the evaluator. Some in the field think that advocacy diminishes the credibility of the evaluator and the evaluation (Chelimsky, 1998; Chen, 1994; Patton, 1997b; Scriven, 1971). Others argue that certain forms of advocacy are important; for example, advocating for democracy and public interest (House & Howe, 1998). Yet, it is clear from our data that evaluators are uncomfortable with idea of being a program advocate.

Related to the issue of the evaluator's role, the results of the logistic regression showed that external evaluators were less likely to be in the "high endorser" group for the planning for evaluation use and organizational outcomes factors. The strengths and limitations of being an external evaluator have been well described (Fitzpatrick, Sanders, & Worthen, 2003; Osborne Daponte, 2008; Russ-Eft & Preskill, 2001), and there is general agreement that external evaluators are often perceived to be more independent and therefore more objective and less biased than internal evaluators. Although there are certain instances where this generalization may not hold, it is useful for us to highlight here because our findings suggest that external evaluators are not completely comfortable implementing activities that nurture organizational learning and change. This may be because doing so could threaten the external evaluators perceived independence and objectivity. More conventional beliefs about internal

evaluators and evaluation use may also help explain our results, for example, because the internal evaluator does not leave the organization at the completion of the evaluation, there is greater opportunity for continued instrumental and process use. Extending this argument would suggest that internal evaluators would also have increased occasion and support for facilitating organizational learning. This would help explain the finding that external evaluators were less likely to be in the “high endorser” group for the planning for use factor.

Another statistically significant result of the logistic regression was that Evaluation Use TIG members were more likely to be “high endorsers” for four of the five factors compared with non-TIG members. This finding was certainly not surprising and is consistent with the descriptive statistics as well. Perhaps more interesting is that this finding substantiates the existence of a cohesive set of opinions and attitudes within the Evaluation Use TIG. Shulha and Cousins (1997) once called the Evaluation Use TIG “a conceptual home” for evaluation theorists, researchers, and practitioners interested in the topic of evaluation use (p. 196). Our data certainly suggest that indeed it does serve this important purpose.

An increased stream of scholarship examining the relationship between evaluative inquiry and organizational learning has emerged since the Preskill and Caracelli 1996 use survey (Cousins et al., 2004; Earl & Katz, 2002; Owen & Rogers, 1999; Preskill & Torres, 1999; Preskill, Zuckerman, & Matthews, 2003; Russ-Eft & Preskill, 2001; Shulha, 2000; Taut, 2007). In 2007, the annual AEA conference theme was evaluation and learning. The presidential address was entitled “Evaluation’s Second Act: A Spotlight on Learning” (Preskill, 2008). Clearly, the focus on organizational learning and other organizational outcomes is expanding and its relevance is growing. In the current study, a majority of respondents endorsed the idea that evaluation fosters organizational outcomes and facilitating organizational learning was the second most agreed upon role of the evaluator.

Study Limitations and Implications

The first limitation relates to the component of our study that compares the 1996 and 2006 study results. The differences in sampling (Use TIG members in 1996 / U.S. members of AEA in 2006), measurement (the survey instruments although similar, did differ), and survey administration method (pen and paper in 1996 / online in 2006) could explain the differences in responses between 1996 and 2006 rather than true changes over time. The second limitation is that the majority of items ask about respondent attitudes or perceptions, not behaviors. High endorsement of an item does not necessarily translate into a corresponding behavior, so we cannot assume that our survey responses reflect actual evaluation practice. The third limitation is that our sample only included U.S. members of the AEA. Datta (2001) argues “estimating the population of evaluators is difficult. Apparently many evaluators in the specialized areas, such as public health, early childhood education, energy, and natural resources, are not members of the American Evaluation Association” (p. 404). Thus, we cannot generalize our findings beyond the attitudes and perceptions of evaluators whom are AEA members and residing in the United States.

This study has implications for practitioners as well as evaluation researchers and theorists. From this study, practitioners gain insight into the attitudes and opinions of other practitioners, which act as a point of comparison for one’s own beliefs, thereby encouraging personal reflection, which is key to forming notions of identity. Descriptive studies such as this one also help draw an outline of the field and those who practice evaluation. It is our hope that the results of this study will also inform future research and provide empirical data to support the development of evaluation theory.

References

- Alkin, M. C. (1985). *A guide for evaluation decision makers*. Beverly Hills, CA: Sage.
- Alkin, M. C. (2003). Evaluation theory and practice: Insights and new directions. *New Directions for Evaluation*, 97, 81-89.
- Alkin, M. C., & Coyle, K. (1988). Thoughts on evaluation misutilization. *Studies in Educational Evaluation*, 14, 331-340.
- Alkin, M. C., Daillak, R., & White, P. (1979). *Using evaluations: Does evaluation make a difference?* Beverly Hills, CA: Sage.
- Alkin, M. C., & Taut, S. M. (2003). Unbundling evaluation use. *Studies in Educational Evaluation*, 29, 1-12.
- American Evaluation Association. (2008, April). *American Evaluation Association Internal Scan Report to the Membership*, by Goodman Research Group. Retrieved September 15, 2008, from www.eval.org.
- Amo, C., & Cousins, J. B. (2007). Going through the process: An examination of the operationalization of process use in empirical research on evaluation. *New Directions for Evaluation*, 116, 5-26.
- Bickman, L., & Reich, S. M. (2009). Randomized controlled trials: A gold standard with feet of clay? In S. I. Donaldson, C. A. Christie, & M. M. Mark (Eds.), *What counts as credible evidence in applied research and evaluation practice?* Thousand Oaks, CA: Sage.
- Chelimsky, E. (1998). The role of experience in formulating theories of evaluation practice. *American Journal of Evaluation*, 19, 35-55.
- Chen, H. (1994). Current trends and future directions in program evaluation. *American Journal of Evaluation*, 15, 229-238.
- Christie, C. A. (2003). What guides evaluation? A study of how evaluation practice maps onto evaluation theory. *New Directions for Evaluation*, 97, 7-35.
- Christie, C. A., & Alkin, M. C. (1999). Further reflections on evaluation misutilization. *Studies in Educational Evaluation*, 25, 1-10.
- Cousins, J. B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan, & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 245-266). Netherlands: Kluwer.
- Cousins, J. B., Goh, S. C., Clark, S., & Lee, L. E. (2004). Integrating evaluation inquiry into the organizational culture: A review and synthesis of the knowledge base. *Canadian Journal of Program Evaluation*, 19, 99-141.
- Cousins, J. B., Goh, S. C., & Elliot, C. (2007, November). *Integrating consequences of evaluation capacity building inquiry*. Paper presented at the annual meeting of the American Evaluation Association, Baltimore, MD.
- Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, 56, 331-364.
- Datta, L. E. (2001). Coming attractions. *American Journal of Evaluation*, 22, 403-408.
- Earl, L., & Katz, S. (2002). Leading schools in a data rich world. In K. Leithwood, P. Hallinger, G. Furman, P. Grann, B. Mulford, & K. Riley (Eds.), *The international handbook for leadership*. Netherlands: Kluwer.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2003). *Program evaluation: Alternative approaches and practical guidelines* (3rd ed.). Boston: Pearson Education, Inc.
- Gersten, R., & Hitchcock, J. (2009). What is credible evidence in education? The role of the What Works Clearinghouse in informing the process. In S. I. Donaldson, C. A. Christie, & M. M. Mark (Eds.), *What counts as credible evidence in applied research and evaluation practice?* Thousand Oaks, CA: Sage.
- Greene, J. G. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review*, 12, 91-116.
- Henry, G. T. (2000). Why not use? *New Directions for Evaluation*, 88, 85-97.
- Henry, G. T. (2009). When getting it right matters: The case for high-quality policy and program impact evaluations. In S. I. Donaldson, C. A. Christie, & M. M. Mark (Eds.), *What counts as credible evidence in applied research and evaluation practice?* Thousand Oaks, CA: Sage.
- Henry, G. T., & Rog, D. J. (1998). A realist theory and analysis of utilization. *New Directions for Evaluation*, 78, 89-102.
- House, E. R. (1980). *Evaluation with validity*. Beverly Hill, CA: Sage.
- House, E. R., & Howe, K. R. (1998). The issue of advocacy in evaluations. *American Journal of Evaluation*, 19, 233-236.
- Julnes, G., & Rog, D. (2009). Evaluation methods for producing actionable evidence: Contextual influences on adequacy and appropriateness of method choice. In S. I. Donaldson, C. A. Christie, & M. M. Mark (Eds.), *What counts as credible evidence in applied research and evaluation practice?* Thousand Oaks, CA: Sage.
- Kirkhart, K. E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation*, 88, 5-23.
- Leviton, L. C., & Hughes, E. F. X. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review*, 5, 525-548.

- Osborne Daponte, B. (2008). *Evaluation essentials: Methods for conducting sound research*. San Francisco: Jossey-Bass.
- Owen, J. M., & Rogers, P. (1999). *Program evaluation: Forms and approaches*. Thousand Oaks, CA: Sage.
- Patton, M. Q. (1994). Developmental evaluation. *Evaluation Practice, 15*, 311-319.
- Patton, M. Q. (1997a). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (1997b). Of vacuum cleaners and toolboxes: A response to Fetterman's Response. *Evaluation Practice, 18*, 267-270.
- Patton, M. Q. (2003). Utilization-focused evaluation. In T. Kellaghan, & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 223-244). Netherlands: Kluwer.
- 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*. Lexington, MA: Heath.
- Preskill, H. (2008). Evaluation's second act: A spotlight on learning. *American Journal of Evaluation, 29*, 127-138.
- Preskill, H., & Caracelli, V. (1997). Current and developing conceptions of use: Evaluation use TIG survey results. *Evaluation Practice, 18*, 209-225.
- Preskill, H., & Torres, R. (1999). *Evaluative inquiry for learning in organizations*. Thousand Oaks, CA: Sage.
- Preskill, H., & Torres, R. (2000). The learning dimension of evaluation use. *New Directions for Evaluation, 88*, 25-37.
- Preskill, H., Zuckerman, B., & Matthews, B. (2003). An exploratory study of process use: Findings and implications for future research. *American Journal of Evaluation, 24*, 423-442.
- Reineke, R. A. (1991). Stakeholder involvement in evaluation: Suggestions for practice. *American Journal of Evaluation, 12*, 39-44.
- Rich, R. F. (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.
- Rossmann, G. B., & Rallis, S. F. (2000). Critical inquiry and use as action. *New Directions for Evaluation, 88*, 55-69.
- Russ-Eft, D., & Preskill, H. (2001). *Evaluation in organizations: A systematic approach to enhancing learning, performance, and change*. New York: Perseus.
- Scriven, M. (1971). Objectivity and subjectivity in educational research. In G. F. Madaus, M. Scriven, & D. L. Stufflebeam (Eds.), *Philosophical redirection of educational research*. Boston: Kluwer-Nijhoff.
- Shadish, W. R., Cook, T. D., & Leviton, L. C. L. (1991). *Foundations of program evaluation*. Newbury Park, CA: Sage.
- Shulha, L. M. (2000). Evaluative inquiry in university-school professional learning partnerships. In V. J. Caracelli, & H. Preskill (Eds.), *The expanding scope of evaluation use. New Directions for Evaluation* (Vol. 88, pp. 39-53).
- Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: Theory, research, and practice since 1986. *Evaluation Practice, 18*, 195-208.
- Smith, N. C. (1993). Improving evaluation theory through the empirical study of evaluation practice. *Evaluation Practice, 14*, 237-242.
- Stufflebeam, D. L., & Shinkfield, A. J. (1985). *Systematic evaluation*. Boston: Kluwer-Nijhoff.
- Taut, S. (2007). Studying self-evaluation capacity building in a large international development organization. *American Journal of Evaluation, 28*, 45-99.
- Taut, S., & Alkin, M. C. (2003). Program staff perceptions of barriers to evaluating implementation. *American Journal of Evaluation, 24*, 213-226.
- Turnbull, B. (1999). The mediating effect of participation efficacy on evaluation use. *Evaluation and Program Planning, 22*, 131-140.
- Weiss, C. H. (1972). Utilization of evaluation: Toward comparative study. In C. H. Weiss (Ed.), *Evaluating action programs: Readings in social action and education*. Boston: Allyn & Bacon.
- Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review, 39*, 426-431.
- Weiss, C. H. (1980). Knowledge creep and decision accretion. *Knowledge: Creation, Utilization, Diffusion, 1*, 381-404.
- Weiss, C. H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation, 19*, 21-33.
- Worthen, B. R. & Sanders, J. R. (Eds.), (1973). *Educational evaluation: Theory and practice*. Worthington, OH: Charles A. Jones.