UCLA
UCLA Previously Published Works

Title
Definitions of Evaluation Use and Misuse, Evaluation Influence, and Factors Affecting Use

Permalink
https://escholarship.org/uc/item/7gw2z6vr

Journal
American Journal of Evaluation, 38(3)

ISSN
1098-2140

Authors
Alkin, Marvin C
King, Jean A

Publication Date
2017-09-01

DOI
10.1177/1098214017717015

Peer reviewed

Definitions of Evaluation Use and Misuse, Evaluation Influence, and Factors Affecting Use

Marvin C. Alkin[1] and Jean A. King[2]

Abstract

The second article in this series on the history of evaluation use has three sections. The first and longest develops a functional definition of the term *use*, noting that a thorough definition of evaluation use includes the initial stimulus (i.e., evaluation findings or process), the user, the way people use the information, the aspect of the program considered, and the purpose. It then defines evaluation use's unethical companion, misuse, detailing the distinction between the two. The second section briefly discusses a broadened concept of evaluation impact that expands to include evaluation influence. Finally, the third section summarizes the factors that research has shown to be related to evaluation use.

Preface

This article is the second of three tracing the development of the concept of evaluation use from the perspective of two academics who have studied the topic for over 40 years. As a reminder, the first article in the series traced the origins and conceptual grounding of evaluation use. Our focus in this second article is on our views of use/misuse and the multiple perspectives that have influenced them over time. We also discuss the addition of evaluation influence to the conceptual debate and factors that are empirically related to use. The third and final article in the series will address evaluation use theories and research.

Evaluation Use and Misuse Defined

The first article in this series (Alkin & King, 2016) discussed "evaluation use," but without a formal definition of the term, although it may be partially inferred by the nature of the earlier discussion. In

[1] University of California, Los Angeles, Los Angeles, CA, USA
[2] University of Minnesota, Minneapolis, MN, USA

Corresponding Author:
Marvin C. Alkin, University of California, Los Angeles, Box 951521, Los Angeles, CA 90095, USA.
Email: alkin@gseis.ucla.edu

considering evaluation use, there seems to be an indication that evaluation information must have consequences of some sort—hence, use. But what exactly do we mean by *use*, and what about its negative relative, *misuse*? These questions are not easily answered. Although researchers appear over time to agree that both use and misuse of evaluation occur, there are not yet consensual definitions.

Background

Let us first—and unavoidably—note that the term *evaluation* means different things to different people (King & Stevahn, 2013, p. 12), but we are not focusing on this distinction, assuming an evaluation process and outcomes as a given. We also assume that the evaluation process is of reasonable quality, that is, that the process is technically adequate and credible and that its findings warrant using. What about situations where the adequacy of an evaluation is called into question? Here we do not refer to situations in which minor technical difficulties occur, but rather where the degree of an evaluation's inadequacy raises major concerns about the validity of the evaluation process and its results. We refer to these poor quality evaluations as "misevaluation." In some instances, what seems to be misuse may actually be attributable to the evaluator's ineptness. As Cousins (2004) notes, "[M]isevaluation can arise from a host of sources including evaluator incompetence, evaluator mischief, unforeseen circumstances leading to design degradation, and the like" (p. 393). We see manifestations of misevaluation

> . . . when the technical aspects of the evaluation have not been conducted adequately (e.g., data collection was done poorly or statistical analyses are incorrect) . . . or when the evaluator fails to understand the evaluation context properly and therefore misdirects the evaluation, . . . or when the evaluator fails to recognize properly his or her obligations for appropriate communication to potential users. (Alkin, 1990, p. 290)

Included within this list may be settings where the evaluator asks the wrong question (Posavac, 1994). Evaluators may also demonstrate poor work when they employ inappropriate methods leading to flawed data collection techniques or failing to qualify the findings. Much like malpractice in the field of medicine, each of these is reflective of bad practice, and thus misevaluation has occurred. A potential user should never participate in a flawed evaluation process, for example, a pseudo-evaluation in which the results are predetermined, or use the results of a technically inadequate study. Years ago, Davis and Salasin (1975) noted that when evaluations are methodologically unsound and "decisions to terminate or reduce programs have been justified on the basis of those evaluations, it is a clear lesson that concern over utilization must emphasize *appropriate* use rather than just greater use" (p. 622).

In addition to this assumption of technical adequacy, we assume that evaluators adhere to the ethical codes of evaluation practice, applying the American Evaluation Association's (AEA, 2004) Guiding Principles, meeting the Program Evaluation Standards (Yarbrough, Shulha, Hopson, & Caruthers, 2010), and attending to the AEA Statement on Cultural Competence in Evaluation (2011). Not surprisingly, these statements have a great deal of overlapping content, and each can provide both implicit and explicit guidance for evaluators seeking to foster evaluation use. They suggest that unethical behavior can take many forms, for example, when evaluators fall prey to requests from program managers to falsify or to modify negative findings, when they change findings in an effort to better their own standing with the program, or when they let strong sympathies for the intent of the program and whom it serves influence decisions. Such behavior is never acceptable and is a particularly disturbing form of misevaluation. We shall return to these points later, but they serve as important grounding for our discussion.

Arriving at consensual definitions for key terms, especially two that have survived for decades without them, is not as simple as it may appear. As Patton (1988, p. 304) noted, citing Samuel Butler, "Definitions are a kind of scratching and generally leave a sore place more sore than it was before." In an effort to avoid additional soreness, let us begin by noting that use and misuse are separate dimensions and therefore better understood if investigated as two separate continua. The use continuum extends from nonuse to use and reflects the extent to which someone does something with an evaluation, although measuring the extent of use may present a nontrivial challenge. Generally, we only seek evidence of whether or not *any* use has occurred as an outcome of an evaluator's efforts. Similarly, the misuse continuum extends from nonmisuse to misuse, meaning that misuse necessarily reflects an ethical dimension, measuring an inappropriate or unethical *manner* of use. As with use, judgment of the extent of misuse along a continuum is difficult and, in an important sense, not necessary; either someone misuses an evaluation or not.

As we shall discuss, in both cases it is important to consider the user's intention as both the evaluation use and misuse continua have the potential to be active or passive. *Active* in this case means that they require a user's purposeful and intentional engagement, what King and Pechman (1982) call "charged use." Use is typically the active extreme of the use continuum; misuse is typically the active extreme of the misuse continuum. Nonuse may, ironically, be "active" (e.g., a user who chooses to ignore the results of a study). Similarly, nonmisuse—an awkward term at best—may also be active (e.g., a user assiduously avoids the potential abuse of evaluation information). Both of these, however, may also be passive (e.g., when a user is not even aware that a study occurred). The following sections will present our perspective on the development of these ideas over time. While it is true that many early writers on evaluation (e.g., Michael Scriven and Daniel Stufflebeam) unavoidably wrote about issues related to use, we focus here on scholars who either wrote extensively about or conducted research on use.

Evaluation Use

Our first article outlined different forms of use, including instrumental, conceptual, and symbolic, and documented that the dispute about the prevalence of use is largely related to the nature and extent of actions taken or thoughts crystallized as a consequence of an evaluation. It also noted that instances of this shaky concept labeled *use* may be derived from evaluation findings as well as from the process, that is, the actions involved in engaging in an evaluation.

Since the earliest discussions of evaluation use, most evaluation writers have not provided a precise definition of the term, but instead commented on the components of what counts as use.[1] Braskamp (1982), for example, avoids a precise definition and says "use cannot singularly be defined; there are four major types of use. Use of an evaluation occurs when any of the following conditions are present." He then stipulates "allocative, direct use" (instrumental), "conceptual enlightenment" (conceptual), "impact on organizational policies and practices as a result of the threat of an evaluation" (symbolic), and "contribution to the management of the organization" (organizational learning or capacity building). He notes that this expanded definition of Pelz's work (1978) incorporates both individual and organizational uses.[2]

Weiss (1990) highlights the potential definition of use as related to conceptual as well as instrumental use. She notes:

> . . . direct and immediate application of evaluation findings to program decisions is not the only kind of use . . . . [T]hese findings provide news; they alert people to new ideas; they show alternative ways of thinking about problems; they alter what is taken for granted as inevitable and what is seen as subject to

change. Evaluation influences people's understanding of what the program has been doing . . . In effect, it can help to change what people think about and how they think about it. This kind of *conceptual* use is an important contribution which, in the long run, can lead—and has led—to significant program reform. (p. 22)

This framing is similar to the criteria Patton et al. (1977) apply, which emphasize direct actions, attitude change, and increase in the forcefulness of an argument.
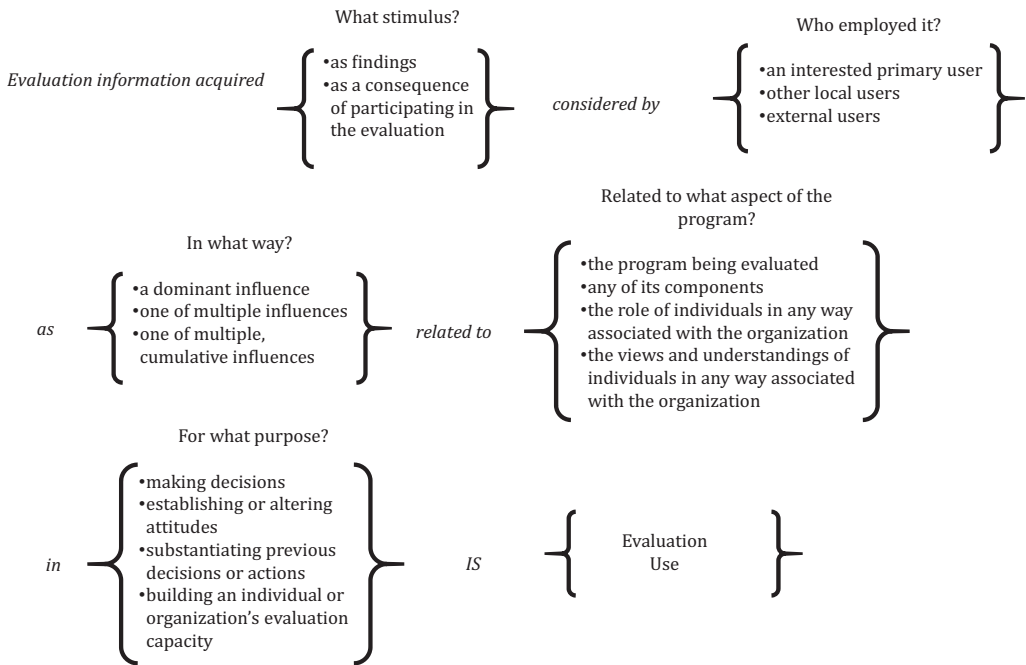
Further, Braskamp and Brown (1980) note that "effective utilization does not necessarily mean that any of the recommendations are implemented or that there are any immediately apparent decisions based on the information" (p. 81). They write that "[m]aking relevant audiences more aware of the program, its context, and the constraints under which it was funded, implemented, and amenable to alterations is regarded as an important end result" (p. viii). However, Leviton and Hughes (1981) talk about "bottom line criteria for utilization" as a start for a definition. They refer to Cook and Pollard's definition (1977), which focused on having a *serious discussion* of findings as they might be related to the program being evaluated—that is, translated into implications for program action. Cousins and Leithwood concur, indicating that use involves the "mere psychological processing of evaluation results . . . without necessarily informing decisions, dictating actions, or changing thinking" (1986, p. 332). Hendricks (1990) takes this idea a step farther: "Our evaluations are utilized when the findings and recommendations from the evaluation are *considered seriously* by persons in a position to act on our information if they choose to do so" (p. 23, emphasis added). Stake writes that "[d]ecision making and policy setting are importance consequences of evaluation study, but *understanding* alone is a consequence that indicates utility in evaluation studies" (1986, p. 141). Leviton and Hughes add an additional criterion: "to be considered used there must be evidence that in the absence of the research (sic) information those engaged in policy or program activities would have thought or acted differently" (1981, p. 527). This, of course, would be extremely difficult to determine.

In 1979, Alkin, Daillak, and White attempted to reconcile the then prevalent and seemingly disparate definitions. Their definition took the form of an adapted Guttman-mapping sentence, a format helpful as a theoretical depiction. Almost 40 years later and based on what we now know, we provide a definition of evaluation use in this format (see Figure 1) and examine the five "matrices" of the Guttman-mapping sentence that define instances of use.

Stimulus. The first matrix considers the stimulus that leads to potential use. The potential for use may stem from either the evaluation findings or its process. The findings (or results) are the summary evaluation information provided in some type of report (e.g., midterm, final; oral, written, enacted). Alternatively, the stimulus may consist of the information or understandings derived as a consequence of participating in the evaluation.

User. The words "considered by" precede the second matrix. In essence, this relates to the previous discussion about people, that is, potential users, seriously considering evaluation information or processes. It is easy to understand how this relates to findings, but more difficult to contemplate when considering the implications related to processes. Clearly, these are more difficult to ascertain, but can possibly be recognized in individual reflections, conversations, and actions.

The content of the second matrix indicates *who* responds to the stimuli. A variety of potential stakeholders may examine the findings or process information. Of key concern are interested primary users, the stakeholders most closely associated with the program who are potentially in a position to make decisions based upon the information. Aside from primary users, other local program users may be in a position to influence decisions, actions, or understandings. Consider, for example, other personnel charged with responsibility for operating a program. We refer to these

What stimulus?

Evaluation information acquired

- as findings
- as a consequence of participating in the evaluation

considered by

Who employed it?

- an interested primary user
- other local users
- external users

In what way?

as

- a dominant influence
- one of multiple influences
- one of multiple, cumulative influences

related to

Related to what aspect of the program?

- the program being evaluated
- any of its components
- the role of individuals in any way associated with the organization
- the views and understandings of individuals in any way associated with the organization

For what purpose?

in

- making decisions
- establishing or altering attitudes
- substantiating previous decisions or actions
- building an individual or organization's evaluation capacity

IS

Evaluation Use

Figure 1. Use defined with an adapted Guttman-scale mapping sentence.

individuals as "other local users." Beyond these individuals or groups, there are additional potential users of the evaluation information or process. Think here about the larger organization that encompasses the program. Think also about staff at agencies who may have funded or supported the program. Consider also the community at large or segments thereof who may be aware of and have interest in the program. Finally, there are the clients of the program. All of these individuals are potentially in a position to respond to the evaluation stimuli. Greene (1988) highlights the importance of stakeholder involvement, distinguishing among the "very involved person," "the somewhat or sometimes involved person," and the "marginally involved person." The extent and way that they may influence potential use determine the extent to which they are included as part of this definition.

Type of influence. In the third matrix, we suggest three ways in which evaluation may affect a program: as a dominant influence, as one of multiple influences at a given time, or as one of multiple influences over time. Evaluation information of whatever type, findings or process, may be a dominant influence on potential actions or understandings. The clearest example may be explicit instrumental use, for example, where a final report includes recommendations and potential users specifically consider those recommendations.

However, the idealized view of the role of evaluation information as the sole (i.e., one and only) input to a program decision or other use is, in practice, naïve in most instances. Previous discussion on the role of "ordinary knowledge" (Lindblom & Cohen, 1979) and "working knowledge" (Kennedy, 1983) attests to the multiplicity of potential inputs to decisions. In sum, there is never really only one input to action. People do not live in a world devoid of all stimuli save for evaluation—although uninformed evaluators sometimes act under that assumption. Alternatively, evaluation

information may be one of multiple influences at a given time. Thus, results provided by data from statewide testing may concur with those presented in the evaluation report, or decision makers may employ evaluation information along with other nonevaluation inputs, such as program costs or district resources. A third and final possibility is that evaluation may be one of multiple, cumulative influences. For example, the evaluation may substantiate findings from an earlier report and thus be sufficient to stir stakeholders to action.

Program aspect considered. As depicted in the fourth matrix, the information a potential user considers may be related to (1) the program as a total entity, (2) to any of its components, or (3) to the functioning of individuals charged with or (4) those associated with responsibility for working within the program. The first two of these are clear, as may be depicted in the following questions: Is the program working? Are there activities that are either not functioning as intended or not attaining appropriate short-term outcomes? The third and fourth refer to any people associated with the program either directly (e.g., a staff member who adapts program activities) or indirectly (e.g., an influential community member in a position to influence the program) who have the ability to affect a program's functioning.

Purpose. Finally, in the fifth matrix, we note that evaluation information and its potential use may relate to several different actions. Thus, people may use the evaluation information in different ways: (1) making decisions, (2) establishing or altering attitudes, (3) substantiating previous decisions or actions, or (4) building an individual's or an organization's evaluation capacity.

While decision-making is a common purpose, sometimes there is no real decision to be made. In some cases critical decisions (e.g., to establish a program) have already been made, and there may be no pressing need to make any program decisions. But people can use evaluation information to alter attitudes or change the climate of opinion surrounding a program (Alkin, Daillak, & White, 1979). Clearly, there is a variety of issues about which one might make or substantiate decisions or alter attitudes. A fourth purpose, which can either apply to individual people or to the organization as a whole, relates to the explicit development of the capacity to both conduct and use the evaluation process and results (Cousins, Goh, Elliott, & Bourgeois, 2014).

In every case, it is helpful to determine the motivation and intent of potential users. For example, if an evaluation report is only commissioned as a symbolic act, we would not consider that an instance of evaluation use because the purpose was not action-oriented. By contrast, if someone provided the report to open-minded people who sought to reaffirm previous decisions, then it would be considered use if potential modifications might occur.

Taken together, the five matrices in this inclusive Guttman-scale definition allow an explicit definition of instances of evaluation use. Consider two such examples:

1. When a manager needs to decide how to modify one aspect of a program based on evaluation data (i.e., "evaluation information acquired as findings considered by an interested primary user as a dominant influence related to a specific program component for making a decision")
2. When an internal evaluator engages in evaluation capacity building by engaging community members in an evaluation (i.e., "evaluation information acquired as a consequence of participating in the evaluation considered by external users as one of multiple influences related to the program being evaluated in building an organization's evaluation capacity")

Although admittedly cumbersome, this functional definition attends to the uniqueness of each instance of use, emphasizing its context-specific nature. It is highly functional for evaluators in that

it clearly describes what use looks like. It lets you know evaluation use when you see it. Thoughtful evaluators can even draw on the Guttman-mapping sentence as an advance organizer for conducting evaluation, actively considering the definition's five components to foster use. In fact, Patton's utilization-focused evaluation (Patton, 2008), first published in 1978, systematically attends to each component, identifying the primary intended users, detailing their hoped-for primary intended uses, and then interacting with them over the course of the evaluation to ensure (to the extent possible) that the process generates credible information that can result in meaningful use.

Misuse

Beginning in the 1960s, as a greater number of evaluations occurred and attention to evaluation use became more prominent, so did attention to instances of *mis*use. To this point, Donald Campbell offered a stern forecast: "The more any quantitative social indicator is used for social decision making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor" (Campbell, 1988, p. 306). Recently Patton (2015) affirmed: "As use increases, misuse will also increase" (p. 142).

   The notion of misuse, then, is not a recent phenomenon but was mentioned early on by Borgatta (1966), Suchman (1967), and Weiss (1972). In 1973, Mushkin warned that "evaluators have not sufficiently safeguarded their statements from misinterpretation" (p. 34). In 1977, Cook and Pollard discussed four studies in which they identified aspects of "misutilization." King and Pechman (1982) conducted research in a school district that resulted in a complicated chart that included examples of misuse. Later, Alkin and Coyle (1988) developed a matrix depicting seven ways in which people could misuse evaluation findings; Christie and Alkin (1999) subsequently expanded upon them. Further, Stevens and Dial (1994a) edited a *New Directions for Evaluation* volume devoted to the topic of misuse.

   What, then, is misuse? Misuse examines evaluation ethics through a lens that differs from most of the evaluation ethics literature, which focuses on the practices and principles of the evaluator (e.g., Newman & Brown, 1996; Morris, 2008). Here we primarily consider the ethics of our clients, those with the potential to use—or, in this case, misuse—the evaluation. As noted above, misuse constitutes a separate continuum, extending from nonmisuse to misuse. It is unethical and therefore inappropriate use—typically focusing on users who manipulate evaluation findings or the evaluation process for personal reasons. One definition of misuse could be the intentional (and even malicious) manipulation of some aspect of an evaluation (e.g., evaluation results) in order to gain something—position or support, for instance. Stevens and Dial (1994a) concur, offering the following definition: "misuse of evaluation means that an evaluation has been used for the wrong purpose or that the results of an evaluation have been misapplied or used improperly" (p. 3). Misuse may also occur due to a failure to use. However, simple nonuse is not really *mis*use since it is unintentional, a passive activity. Misuse is present only when the nonuse actively occurs because of ethical concerns, such as when stakeholders intentionally disregard evaluation results, not based upon merits or the availability of other persuasive data sources.

   With one change, the Guttman-scale definition of evaluation use works to define evaluation misuse. As was true for use, the categories of stimulus, user, type of influence, and program aspect delimit the context. What differs in instances of misuse is the user's purpose, which in the case of misuse is always unethical, that is, for self-interest and personal gain. Consider two examples of misuse:

1. When an administrator decides to cherry pick the results of a study to keep an ineffective program funded (evaluation findings considered by an interested primary user as a dominant influence related to the program being evaluated in promoting a program that is ineffective)
2. When staff members take part in an evaluation process to highlight negative data about a program activity they dislike (evaluation process manipulated by a local user as one of multiple influences related to a program component)

Table 1. Types of Evaluation Misuse.

Category of Misuse	Form of Misuse
	When the evaluation process is good
Commissioning	User commissions evaluation for political show or as a delaying action (symbolic use)
	User commissions evaluation to justify decisions already made and not open to possible changes (legitimative use)
Process	User subverts the evaluation process (e.g., by cutting funding for it, by limiting access to data sources or existing data)
Findings	User modifies data or report content or cherry picks content intentionally
	User actively misrepresents evaluation information (e.g., actively changing it, distributing incomplete results)
	User purposefully ignores evaluation findings for personal gain
	User actively uses evaluation information known to be inaccurate or invalid

Over the years, scholars have documented that misuse can take place at various stages of the evaluation engagement, including: (1) when *commissioning* an evaluation; (2) during the *evaluation process* itself; or (3) when dealing with the *evaluation findings* (Christie & Alkin, 1999). Misuse can also occur when someone chooses to use the outcomes of a poorly conducted evaluation study (i.e., misevaluation). Let us examine each of these. Table 1 summarizes these possibilities.

Commissioning misuse. Evaluations initiated or commissioned for any purpose other than to help inform programmatic decisions are examples of commissioning misuse. Implicit in these actions is the user's deliberate intention not to actually use the evaluation's results or process. Upon commissioning the evaluation, users may, for example, have previously determined decisions about a program or have no real willingness or intention to use the evaluation findings.

As shown in Table 1, two common types of use lend themselves to potential misuse. In fact, we believe that symbolic "use" is not ultimately a type of use; it is misuse. Numerous authors have commented on the commissioning of an evaluation for symbolic reasons, for example, for political gain or for publicity (Weiss, 1973) or merely to gain funding (Duffy, 1994; Stevens & Dial, 1994b). Commissioners may sponsor evaluations only for political purposes, for example, to satisfy funding sources with no real intention of paying attention to the evaluation or its use. Another political purpose may be to place oneself into the public eye such as to enhance one's professional prestige (Alkin et al., 1979). Results from such evaluations are rendered meaningless. This is, in essence, the use type previously referred to as symbolic use, which, owing to its unethical commitment and its potential waste of organizational resources, is actually a form of misuse.

There is another form of symbolic use. Misuse may occur on the part of stakeholders and other potential users by purposefully commissioning the evaluation in order to delay actions or to avoid taking responsibility with no intent to use the results (Suchman, 1967; Weiss, 1973). In this instance, potential users may be faced with a decision and perhaps choose not to deal with it. They may also view commissioning an evaluation as a holding action to avoid proceeding. The results of the evaluation are not relevant; it is the conduct of the evaluation ritual that is relevant. This, too, would constitute misuse.

Another common scenario, which Owen (2002) labels legitimative use, occurs when people commission evaluations to justify decisions already made. Frequently, but not always, legitimative

use is misuse. If the intent is only to justify a previous decision with little or no possibility that evaluation results will lead to a reconsideration of previous decisions, that, too, constitutes misuse. However, if the legitimative use is to confirm the decision or may have, as a consequence, the possible modification of the decision or action, that legitimative use is not misuse.

Evaluation process misuse. Table 1 also shows that misuse may occur during the course of the evaluation's conduct. Administrator actions during the evaluation may seek to delay critical decisions (Christie & Alkin, 1999). King (1988) has commented with respect to process misuse that decision-makers can use political influence to subvert or sabotage an evaluation by not supporting various elements of the evaluation process, resulting in an incomplete evaluation. This is often the case when administrators anticipate findings they perceive or fear to be unsatisfactory. People may see unsatisfactory results as potentially jeopardizing the life of a program, the reputation of an agency, or the political standing of an administrator. Subverting part of an evaluation during its conduct could make the evaluation less meaningful and its credibility subject to attack. Process misuse may also occur related to the financial aspects of conducting an evaluation. For example, decision-makers may divert funds intended to support the evaluation to support other efforts (Christie & Alkin, 1999). These funds may go to increasing program support or, as we have sometimes seen, even be directed to other programs.

Findings misuse. The misuse of evaluation findings, likely the most common type of misuse, can take many forms. Again, we note that proper instances of use do not require that findings or recommendations (if any) be totally heeded. Rather, the issue is that potential users *consider* evaluation findings as part of a decision-making process. Misuse of findings often occurs because evaluations yield results that users perceive as undesirable or uncomfortable. This form of misuse occurs when evaluation evidence does not fit someone's predetermined agenda. Table 1 includes four common ways in which such findings misuse may occur. First, users may modify or select findings that best suit their agenda. In this case, users take evaluation findings out of context or misinterpret them for their own advantage. Thus, a program director may excerpt a portion of the evaluation findings, omitting the caveats and the qualifying remarks to present a more positive result than the findings support. Chelimsky (2011) provides an example of this in what she refers to as the "single narrative."

More blatantly, a second approach involves results being totally misrepresented. Potential users can exaggerate or inaccurately transmit results (Weiss & Bucuvalas, 1980). In this case, they may present findings that deviate from the actual results (Cook & Pollard, 1977). Alternatively, potential users can reshape and rewrite evaluation conclusions (House, 1980). Another form of misuse occurs when people disseminate incomplete and misleading results before a final evaluation report is complete (Cook & Pollard, 1977), despite admonitions on the part of the evaluator about the tentative nature of the findings. The early release of findings can result in misinformed decision-making.

There are instances where the user purposefully ignores the evaluation findings. This nonuse may (or may not) be misuse. We point out that if there is no ethical dimension to such cases; these are simple instances of nonuse. But, when the user purposefully ignores evaluation findings for personal gain an ethical issue emerges. This, also, is misuse.

Another instance of misuse (or perhaps mistaken use) is based on our understanding that evaluations always occur within a political context (Palumbo, 1994; Patton, 2008; Weiss, 1973). Some instances of actions by users may appear to be misuse but an understanding of context would dismiss that notion. Misuse may occur for perfectly understandable and innocent reasons. For example, if misevaluation has occurred, nonuse of inaccurate findings may be the appropriate option (King, 1988). The intention to use an inadequate evaluation is a function of users' technical knowledge and skills, that is, their ability to discern that misevaluation has taken place. If someone is not technically

competent or aware of the attributes of quality evaluation, we cannot fairly label their use of a poor study misuse; rather it is mistaken use.

By contrast, if the user is informed and methodologically capable of discerning the poor quality of an evaluation—and still uses it that should be categorized as an instance of misuse. In these cases, misuse occurs when a user knowingly chooses to use the inadequate results for personal gain. For example, it is misuse when an administrator knows that a survey is filled with bad items and has an extremely low response rate yet uses its results to justify a decision. One major determinant of misuse in the event of misevaluation, then, has to do with the technical sophistication of the user.

As noted at the beginning of this article, evaluation writers have described types of evaluation use in the absence of a common definition. Our solution to this challenge is to define use with a Guttman-type scale to be clear about the components to include in a functional definition. As described previously, a thorough definition of evaluation use includes five components: (1) the initial stimulus (i.e., evaluation findings or process), (2) the user, (3) the kind of influence the stimulus has, (4) the aspect of the program considered, and (5) the purpose of the use. The definition and discussion of use's undesirable companion, misuse, distinguish between the two by highlighting the purpose that motivates a user in a given political context to act unethically. The next section will discuss a more recent addition to thinking on evaluation use that suggests the value of a broadened concept of evaluation impact.

Thinking About Broader Evaluation Impact: Evaluation Influence

Early on, scholars (e.g., Leviton & Hughes, 1981; Weiss, 1972) recognized the potential for program evaluation to have impact beyond individual programs. However, the bulk of writing and research on evaluation use has focused on specific program evaluations and settings, that is, looking at the direct use or misuse of evaluation results or processes in context. An expanded notion of evaluation use early in the new millennium re-conceived the field's attempts to understand its impact (King & Stevahn, 2013). Almost two decades ago, Kirkhart introduced the term *evaluation influence* as an expansion and putative improvement to the long-standing concepts of evaluation use and misuse, highlighting the potential of thinking more broadly about the effects and consequences of evaluation. She contended that the designation *evaluation use* was too limiting, believing that evaluations have impact in many ways, in many places, and at many times. Evaluation influence is "the capacity or power of persons or things to produce effects on others by intangible or indirect means" (Kirkhart, 2000, p. 7). Kirkhart argued that to examine how evaluation affects and changes society writ large, scholars should step back from a narrow focus on use in favor of this broader-based construct. She proposed keeping the term *use* for examples that were specific, but adding the term *influence* as a better term for fostering a more "inclusive understanding of the impact of evaluations" (Kirkhart, 2000, p. 5). As Kirkhart proposed it, "Encapsulating existing insights and approaches, evaluation influence offers a comprehensive framework with which to consider the intended and unintended impacts that evaluation can have . . ." (Herbert, 2014, p. 393).

As originally conceived, evaluation influence focused on the *in*direct, *in*tangible influence that evaluation studies can have on individuals, programs, communities, and systems. The intent was to expand possibilities beyond the focus on direct use of evaluation results or processes. Kirkhart (2000) outlined an "integrated theory of evaluation influence" that included three key variables—source (process/results), intention (intended/unintended), and time (immediate/end of cycle/long-term). Applying *evaluation influence* as a unifying construct, research could then lead to a more thorough, integrated understanding of evaluation's consequences over time. Kirkhart summarized, "This integrated theory of influence helps us recognize that evaluation practice has had a more pervasive impact than heretofore perceived" (p. 20). "Unbundling" the concept of evaluation influence, Alkin and Taut (2003) suggested a revision of Kirkhart's "theory" that added *awareness* as a

variable and then proposed that evaluators address immediate and end-of-cycle use of both the evaluation process and its results. They removed influence—which is linked to unintended outcomes over which evaluators have no control and of which they are unaware—from evaluators' concerns.

In two often cited articles, Henry and Mark (2003) and Mark and Henry (2004) outlined ways to "move the field beyond use and toward a focus on evaluation influence" (Henry & Mark, 2003, p. 295), detailing how Kirkhart's concept might advance research about what occurs when evaluations enter multiple systems. For Henry and Mark, evaluation is about sense making, and it can take different pathways en route to social betterment. Noting that "important shortcomings still exist in previous work on use" (Mark & Henry, 2004, p. 35), they argued that understandings of evaluation use were simultaneously "overgrown" and "underdeveloped." These understandings were over-grown owing to overlapping forms of use, ambiguous constructs, and a lack of indicators, among other features. They were "underdeveloped" because "current models of use are generally silent on the range of *underlying mechanisms* through which evaluation may have its effects" (Mark & Henry, 2004, p. 37). In their thinking, "influence, combined with the set of mechanisms and interim outcomes, 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" (Henry & Mark, 2003, p. 293).

Applying a realist evaluation approach, they proposed a model of "alternative mechanisms that may mediate evaluation influence" (general, cognitive and affective, motivational, and behavioral) at three levels (individual, interpersonal, and collective). Then, expanding on the logic model of evaluation (Cousins, 2003), they also proposed a "schematic theory of evaluation influence," which included evaluation inputs, activities, outputs, general mechanisms, and intermediate and long-term outcomes, all in the grounding context of environmental contingencies and leading ultimately to social betterment.

Nunneley, King, Johnson, and Pejsa (2015) detailed three problems inherent in the Henry and Mark arguments that detract from this theory's potential. First, it provides only imprecise and vague definitions of influence. Second, the schematic theory lacks the attributes of a strong theory (e.g., utility in prediction and explanation). Third—and ironically—the discussion of influence contains a suppressed premise in logic that necessarily requires evaluation use, broadly defined, for influence to occur. In cases of influence, someone must respond in some way to some process or artifact from an evaluation (see Figure 1), that is, they must *use* the evaluation, making use a requisite part of influence rather than distinct from it.

What, then, is the potential value of a broadened concept of evaluation impact that includes evaluation influence? Is influence a helpful addition to the concept of evaluation use? To us the answer is yes in that it has affirmatively stimulated people in the field to think about the wide-ranging effects of evaluation, especially in terms of broader systems thinking, which is clearly an important step. We do not believe, however, that a complete shift in terminology from use to influence is necessary, a point that, for space reasons, we will discuss further in the third article in this series.

Factors Associated With Evaluation Use

In setting out a case for research on evaluation use, Weiss (1972) established an agenda for research on factors associated with evaluation. Many areas that Weiss suggested for future inquiry proved to be major findings of subsequent research as researchers in the 1970s and 1980s were stimulated to consider evaluation use. Teams of evaluation scholars around the country conducted important early studies of factors associated with use: Alkin et al. (1979), Braskamp, Brown, and Newman (1978; Brown, Braskamp, & Newman, 1978), King, Pechman, and Thompson (King & Pechman, 1984; King & Thompson, 1983), and Patton and colleagues (1977). As studies from a variety of

perspectives began to accumulate, evaluation scholars sought to bring cohesion and practical value to them by compiling lists of factors affecting use that they had identified.

An article by Leviton and Hughes (1981) is one of the earliest attempts at drawing together the literature on factors associated with evaluation use. This work relied on the authors' perception of the important factors substantiated by writings drawn from both the research and conceptual literature. References were from papers alluding to "research use," as well as "evaluation use." The major categories that Leviton and Hughes identified were the following: relevance of the evaluation to the users' needs, communication between the individuals performing the evaluation and the users of the evaluation, extent to which findings translated into actionable implications, credibility of the evaluation, and user commitment.

Alkin (1985) is another early compendium of studies focused on evaluation use factors. This analysis drew heavily from findings and a major qualitative research study published as a book in the Sage Research Series (Alkin et al., 1979). In addition to referencing his own earlier work, Alkin cited verified research studies of evaluation use as well as references from the conceptual literature. Major categories identified in this compendium included human factors: users; human factors: the evaluator; evaluation activities; and organizational/social/political factors. An especially important element was the commitment to use on the part of both evaluators and key stakeholders. To obtain the latter, the study identified engaging in the training of stakeholders to be users as an important factor.

Over the ensuing years, three major compilations of evaluation use factors appeared in the literature, each of which excluded content from conceptual pieces, focusing only on verifiable research studies:

- Cousins and Leithwood (1986). This thorough analysis of 65 empirical studies of evaluation use framed its review using 12 factors, six grouped as evaluation implementation factors (evaluation quality, credibility, relevance, communication quality, findings, and timeliness) and another six grouped as decision- or policy-setting factors (information needs, decision characteristics, political climate, competing information, personal characteristics, user commitment and/or receptiveness to evaluation). Cousins and Leithwood discussed the results for each factor, giving the percentage of studies on that factor (never more than 40%) and the conclusiveness of the results. They then calculated the relative influence of factors, noting that it "varied as a function of the type of use" (Cousins & Leithwood, 1986, p. 359).
- Shulha and Cousins (1997). This study examined research studies conducted since the Cousins and Leithwood review. It highlighted four changes and developments: (1) the emergence of context as a critical variable in considering use, (2) the identification of process use as potentially influential, (3) the expansion of the concept of use from the individual to the organizational level, and (4) the emerging role of the evaluator as a facilitator and teacher.
- Johnson, Greenseid, Toal, King, Lawrenz, and Volkov (2009). Using the same framing as Cousins and Leithwood (i.e., 12 factors grouped under evaluation implementation and decision-/policy-setting), this study analyzed research on evaluation use factors conducted since the Shulha and Cousins review. The major change since the previous review was an increased emphasis on stakeholder involvement,

  . . . a mechanism that facilitates those aspects of an evaluation's process or setting that lead to greater use. More than just involvement by stakeholders or decision makers alone, however, the findings from this literature review suggest that engagement, interaction, and communication between evaluation clients and evaluators is key to maximizing the use of the evaluation in the long run. (Johnson et al., 2009, p. 389)

In addition to these formal research compilations, the *Program Evaluation Standards* (Yarbrough et al., 2010) are another source of factors grounded in empirical studies. As the introduction to the

Utility Standards puts it, "At its simplest, judgments about an evaluation's utility are made based on the extent to which program stakeholders find evaluation processes and products valuable in meeting their needs" (Yarbrough et al., 2010, p. 4). The third edition's utility standards, eight in all, make explicit what evaluators should consider as they seek to plan and implement useful evaluations: U1 *evaluator credibility*, U2 *attention to stakeholders*, U3 *negotiated purposes*, U4 *explicit values*, U5 *relevant information*, U6 *meaningful processes and products*, U7 *timely and appropriate communicating and reporting*, and U8 *concern for consequences and influence*.

If you examine listings from the three major studies and the *Standards*, you will identify many overlaps. However, tabulating the extent of total overlap present is difficult because the studies sometimes employ different terms that seemingly describe the same concept. Johnson et al. (2009) speak to the practical outcome of this situation in summarizing their review of empirical research on evaluation use: "It is impossible, finally, to answer the question of which characteristics are most related to increasing the use of evaluations in a straightforward manner" (Johnson et al., 2009, p. 388). Nevertheless, an analysis of the factors identified enables us to develop themes that have prevailed over time across the body of literature on factors associated with evaluation use.

Applying the commonplaces of evaluation framework (King, 1988), we have categorized these evaluation use factors into four groupings: (1) user factors, (2) evaluator factors, (3) evaluation factors, and (4) organizational/social context factors. In reality, all four of these categories are part of the context of any evaluation. Those looking for practical application of these factors in practice may find Patton (2008, 2012) or King and Stevahn (2013) helpful.

User factors. Patton et al.'s research (1977) at the University of Minnesota focused the evaluation community's attention on users, identifying the "personal factor" as an important element in determining whether people used evaluations. Patton's work and the focus of other researchers defined as a prime element within the "user factors" category the user's attitude towards evaluation, both in general and specific to the current evaluation. This interest is partially derived from the users' meaningful involvement in the evaluation (Johnson et al., 2009). Users' predispositions about evaluation, based in part on positive prior experiences, color their interest in evaluation (Alkin et al., 1979). While there were other factors identified within this category, as in the evaluator characteristics category described next, the importance of an explicit commitment to use predominated.

Evaluator factors. Particularly relevant within the "evaluator factors" category was a dedication and commitment to facilitating and stimulating use on the part of the evaluator. This evaluator personal factor identified by Alkin, Daillak, and White (1979), AKA the interpersonal factor (King & Stevahn, 2013), is comparable to the personal factor for users that Patton noted. Further, evaluators needed to show political sensitivity and, as Utility Standard 1 affirms, have credibility. Likewise of importance was the way in which evaluators involved or engaged potential users. This consisted of both engaging potential users in various aspects of the evaluation's conduct as well as developing rapport and a good working relationship with those users (Johnson et al., 2009). Each of these in turn had impact on the credibility of the evaluator. Indeed, as we will readily note, many of the identified factors have strong interrelationships.

Evaluation factors. This is the third major category. We distinguish the evaluation and the activities within it from the unique role of the evaluator as an individual. Here the literature speaks to such things as evaluation procedures, the relevance of evaluation information, and communication quality, including evaluation reporting (Shulha & Cousins, 1997). For evaluation procedures, the research shows that while technical excellence and necessary rigor are important, what is most relevant is the appropriateness of the methods employed and their credibility with potential users.

The factor named "relevance" refers to the extent to which the information needs of the program are met (King & Pechman, 1982; Yarbrough et al., 2010). If the evaluation information collected does not meet the perceived need of users, then, not surprisingly, they are not likely to use it.

Finally, within this category is the issue of communication quality. In some ways, this is similar to our previous discussion about the appropriateness of methods. Reporting must be in a form that users can understand. Further, the evaluation information to be reported must be timely, or it may be of no use to potential stakeholders (Alkin, Kosecoff, Fitzgibbon, & Seligman, 1974). Another factor is somewhat more complex; of particular significance is the relationship of the evaluation to existing or competing information either generally available or a part of the working knowledge of major users in the program.

Organizational/social context factors. The fourth major category framed within this analysis related to "organizational/social context factors." Here, several of the researched factors recognized that the nature of the organization in which the evaluation is being conducted has substantial impact on the successful achievement of evaluation use. These varied factors included not only organizational characteristics of the program, but of the larger programmatic entity that encompasses that program (Johnson et al., 2009). Also considered here were issues such as the extent of unit level autonomy and various institutional arrangements. Issues about the age of the program and the extent of its development both affected the extent to which people could productively use an evaluation. In addition, various external factors, namely the community, its influence on the program, and the role of other agencies, are within this category (Alkin et al., 1979). But perhaps of most significance were those sources of information beyond the evaluation that are likely to be employed in making a decision (Patton et al., 1977).

While these studies on individual factors do not constitute a full-blown theory of evaluation use, this general understanding about factors associated with use is not an insignificant accomplishment. Indeed, Lincoln and Guba (2004) cogently comment on the work of Abraham Kaplan, who makes the point that there are other kinds of theories, so-called *pattern theories*, that describe how things come to cohere to form recognizable patterns. "These patterns may be behaviors and activities, social or professional constellations, or an organizational ecosystem, for instance. Their hallmark characteristic is not a set of laws that promise prediction, but rather a series of 'tendency statements'" (Lincoln & Guba, 2004, p. 227) that have meaning in their joint application. So, too, we believe, do these various researched factors form a pattern theory of evaluation use. Moreover, as will be discussed in the final article in this series, we believe that this expanded factor listing provides a useful starting point for further research on the topic.

Summary and Conclusion

In concluding this second article, we again note that the practice of evaluation is, at its core, concerned with the utility of both its process and its results, making evaluation use an integral feature of evaluation practice and theorizing. This article presented the developmental background and a functional "definition" applicable to both evaluation use and misuse, the expanded concept of evaluation influence, and four categories of factors that research has linked to evaluation use (factors related to the user, the evaluator, the evaluation process and outcomes, and the organizational/social context). The third and final article in this series will focus on theories of evaluation use, tracing the historical evolution of theories of and research on evaluation use. This will include the prescriptive theories presented on the use branch of Alkin's evaluation theory tree (2013) along with the two theories of evaluation influence: Kirkhart's integrated theory of evaluation influence and Henry and Mark's schematic theory of evaluation influence.

Declaration of Conflicting Interests

Funding

Notes

1. One of the article's reviewers helpfully noted that "conceptual clarity is a desideratum in our field," stating that 100% agreement on definitions may be both unrealistic and unnecessary and identifying the potential value of ongoing critical discussion of key concepts.
2. Weiss, Murphy-Graham, and Birkeland (2005) proposed a quirky addition to the categories of use: imposed use. In this case, an external authority requires potential users to respond in prescribed ways to the results of a highly structured evaluation process with predefined standards. Commenting on these case studies, Patton (2015) posits that so-called imposed use may in some instances be misuse.

References

Alkin, M. C. (1985). *A guide for evaluation decision-makers*. Beverly Hills, CA: Sage.

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

Alkin, M. C. (Ed.). (2013). *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed.). Los Angeles, CA: Sage.

Alkin, M. C., & Coyle, K. (1988). Thoughts on evaluation utilization, misutilization and non-utilization. *Studies in Educational Evaluation*, *14*, 331–340.

Alkin, M. C., Daillak, R., & White, P. (1979). *Using evaluations: Does evaluation make a difference?* (Vol. 76). Sage Library of Social Research. Beverly Hills, CA: Sage.

Alkin, M. C., & King, J. A. (2016). The historical development of evaluation use. *American Journal of Evaluation*, *37*, 568–579.

Alkin, M. C., Kosecoff, J., Fitzgibbon, C., & Seligman, R. (1974). *Evaluation and decision-making: The Title VII experience*. Los Angeles, CA: Center for the Study of Evaluation.

Alkin, M. C., & Taut, S. (2003). Unbundling evaluation use. *Studies in Educational Evaluation*, *29*, 1–12.

American Evaluation Association. (2004). *AEA's guiding principles for evaluators*. Retrieved from http://www.eval.org

American Evaluation Association. (2011). *AEA's statement on cultural competence in evaluation*. Retrieved from http://www.eval.org

Borgatta, E. F. (1966). Research problems in evaluation of health service demonstrations. *The Milbank Memorial Fund Quarterly*, *44*, 182–201.

Braskamp, L. A. (1982). A definition of use. *Studies in Educational Evaluation*, *8*, 169–174.

Braskamp, L. A., & Brown, R. D. (Eds.). (1980). Utilization of evaluation information. *New Directions for Program Evaluation*. San Francisco, CA: Jossey-Bass.

Braskamp, L. A., Brown, R. D., & Newman, D. L. (1978). The credibility of a local educational program evaluation report: Author, source, and client audience characteristics. *American Educational Research Journal*, *15*, 441–450.

Brown, R. D., Braskamp, L. A., & Newman, D. L. (1978). Evaluator credibility as a function of report style: Do jargon and data make a difference? *Evaluation Quarterly*, *2*, 331–341.

Campbell, D. T. (1988). *Methodology and epistemology for social science: Selected papers* (E. S. Overman). Chicago IL: University of Chicago Press.

Chelimsky, E. (1983). Improving the cost effectiveness of evaluation. In M. C. Alkin & L. C. Soloman (Eds.), *The costs of evaluation* (pp. 149–170). Newbury Park, CA: Sage.

Chelimsky, E. (2011). *Evaluation and the single narrative*. Presentation at the American Evaluation Association Conference, Anaheim, CA.

Christie, C. A., & Alkin, M. C. (1999). Further reflections on evaluation misutilization. *Studies in Educational Evaluation*, *25*, 1–10.

Cook, T. D., & Pollard, W. E. (1977). Guidelines: How to recognize and avoid some common problems of misutilization of evaluation research findings. *Evaluation*, *4*, 161–164.

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, the Netherlands: Kluwer Academic Press.

Cousins, J. B. (2004). Commentary: Minimizing evaluation misuse as principled practice. *American Journal of Evaluation*, *25*, 393–399.

Cousins, J. B., Goh, S. C., Elliott, C. J., & Bourgeois, I. (2014). Framing the capacity to do and use evaluation. *New Directions for Evaluation*, *141*, 7–23.

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

Davis, H., & Salasin, S. (1975). The utilization of evaluation. *Handbook of Evaluation Research*, *1*, 621–666.

Duffy, B. P. (1994). Use and abuse of internal evaluation. *New Directions for Program Evaluation*, *64*, 25–32.

Greene, J. C. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review*, *12*, 91–116.

Hendricks, M. (1990). Participant introductory comments. In M. C. Alkin (Ed.), *Debates on evaluation* (p. 23). Newbury Park, CA: Sage.

Henry, G. T., & Mark, M. M. (2003). Beyond use: Understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation*, *24*, 293–314.

Herbert, J. L. (2014). Researching evaluation influence: A review of the literature. *Evaluation Review*, *38*, 388–419.

House, E. R. (1980). *Evaluating with validity*. Beverly Hills, CA: Sage.

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation*, *30*, 377–410.

Kennedy, M. M. (1983). Working knowledge. *Science Communication*, *5*, 193–211.

King, J. A. (1988). Research on evaluation use and its implications for evaluation research and practice. *Studies in Educational Evaluation*, *14*, 285–299.

King, J. A., & Pechman, E. M. (1982). *The process of evaluation use in local school settings* (Final report of National Institute of Education Grant NIE-G-81-0900). New Orleans, LA: New Orleans Public Schools.

King, J. A., & Stevahn, L. (2013). *Interactive evaluation practice: Mastering the interpersonal dynamics of program evaluation*. Los Angeles, CA: Sage.

King, J. A., & Thompson, B. (1983). Research on school use of program evaluation: A literature review and research agenda. *Studies in Educational Evaluation*, *9*, 5–21.

Kirkhart, K. E. (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*, *88*, 5–23.

King, J. A., & Pechman, E. M. (1984). Pinning a wave to the shore: Conceptualizing school evaluation use. *Educational Evaluation and Policy Analysis*, *6*(3), 241–251.

Leviton, L. C., & Hughes, E. F. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review*, *5*, 525–548.

Lincoln, Y. S., & Guba, E. (2004). The roots of fourth generation evaluation: Theoretical and methodological origins. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 225–242). Thousand Oaks, CA: Sage.

Lindblom, C. E., & Cohen, D. K. (1979). *Usable knowledge: Social science and social problem solving* (Vol. 21). New Haven, CT: Yale University Press.

Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, *10*, 35–57.

Morris, M. (Ed.). (2008). *Evaluation ethics for best practice: Cases and commentary*. New York, NY: Guilford Press.

Mushkin, S. J. (1973). Evaluations: Use with caution. *Evaluation*, *1*, 30–35.

Newman, D. L., & Brown, R. D. (1996). *Applied ethics for program evaluation*. Thousand Oaks, CA: Sage.

Nunneley, R. D., King, J. A., Johnson, K., & Pejsa, L. (2015). The value of clear thinking about evaluation theory. In A. Vo (Ed.), *Evaluation use and decision-making in society: A tribute to Marvin C. Alkin* (pp. 53–71). Denver, CO: Information Age.

Owen, J. M. (2002). *Linking evaluation use to the research utilization literature*. Paper presented at the Annual Meeting of the American Evaluation Association, Arlington, VA.

Palumbo, D. J. (1994). The political roots of misuse of evaluation. *New Directions for Program Evaluation*, *64*, 15–23.

Patton, M. Q. (1988). Six honest serving men for evaluation. *Studies in Educational Evaluation*, *14*, 301–330.

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Los Angeles, CA: Sage.

Patton, M. Q. (2012). *Essentials of utilization-focused evaluation*. Los Angeles, CA: Sage.

Patton, M. Q. (2015). Misuse: The shadow side of use. In C. A. Christie & A. T. Vo (Eds.), *Evaluation use and decision making in society: A tribute to Marvin C. Alkin* (pp. 131–147). Charlotte, NC: Information Age.

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. 141–163). New York, NY: D. C. Health.

Pelz, D. C. (1978). Some expanded perspectives on use of social science in public policy. In J. M. Yinger & S. J. Cutler (Eds.), *Major social issues: A multidisciplinary view* (pp. 346–357). New York, NY: Free Press.

Posavac, E. J. (1994). Misusing program evaluation by asking the wrong question. *New Directions for Program Evaluation*, *64*, 70–78.

Rossman, G. B., & Rallis, S. E. (2000). Critical inquiry and use as action. *New Directions for Evaluation*, *88*, 55–69.

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

Stake, R. E. (1986). *Quieting reform: Social science and social action in an urban youth program*. Urbana: University of Illinois Press.

Stevens, C. J., & Dial, M. (Eds.). (1994a). Preventing the misuse of evaluation. *New Directions for Program Evaluation*, *64*, 1–84.

Stevens, C. J., & Dial, M. (1994b). What constitutes misuse? *New Directions for Program Evaluation*, *64*, 3–13.

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

Weiss, C. H. (1972). Utilization of evaluation: Toward comparative study. *Evaluating action programs: Readings in social action and education*, 318–326. Boston, MA: Allyn and Bacon.

Weiss, C. H. (1973). Where politics and evaluation research meet. *Evaluation*, *1*, 37–45.

Weiss, C. H. (1990). Participant introductory comments. In M. C. Alkin (Ed.), *Debates on evaluation* (p. 22). Newbury Park, CA: Sage.

Weiss, C. H., & Bucuvalas, M. J. (1980). Truth tests and utility tests: Decision-makers' frames of reference for social science research. *American Sociological Review*, *45*, 302–313.

Weiss, C. H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence how evaluations affect DARE. *American Journal of Evaluation*, *26*, 12–30.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2010). *The program evaluation standards: A guide for evaluators and evaluation users*. Thousand Oaks, CA: Sage.