The Centrality of Use: Theories of Evaluation Use and Influence and Thoughts on the First 50 Years of Use Research

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

Abstract

Our third article on the history of evaluation use affirms its importance in evaluation practice and related literature. It first highlights the centrality of use in the field's professionalizing documents, extant theories, and the persistence of continuing research. Next, it discusses the challenge of evaluation theories in general, including the prevalence of prescriptive theories, and provides criteria for a good descriptive use theory that even the most detailed use theory does not currently meet. The third section reviews existing use "theories," nine from Alkin's evaluation theory tree, two from the literature, and two existing influence theories. This article concludes with a discussion of research on evaluation use past and present, including the effects of competing definitions, along with thoughts for future inquiry.

Preface

This article, the third in a series documenting the historical record of the concept of evaluation use, details the centrality of use in current practice and examines theories of and research on evaluation use and influence. Our purpose in compiling the history of this evolution is to help evaluators better understand the research grounding of evaluation use, making clear the potential role of specific practices, the related effects of organizational context, and how best to study them.

To review, the first article in the series traced the evolution of this increasingly pervasive concept, documenting two developmental streams (educational testing and measurement and the social sciences), detailing competing perspectives about the prevalence of use, and presenting its traditional categories in the literature. The second article began by exploring definitions of use and its

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

Corresponding Author:
Jean A. King, University of Minnesota, 206 Burton Hall, 178 Pillsbury Drive SE, Minneapolis, MN 55455, USA.
Email: kingx004@umn.edu

opposite, misuse, then described an expanded concept labeled evaluation influence and created a "pattern theory" by summarizing research-grounded factors related to increased use.

As noted in Article 2 (Alkin & King, 2017, p. 447), the distinctive characteristic of a pattern theory "is not a set of laws that promise prediction, but rather a series of 'tendency statements'" (Lincoln & Guba, 2004, p. 227) that have meaning when applied together. The pattern theory of evaluation use suggests that evaluation use results when certain types of users (i.e., those committed to use) interact with certain types of evaluators (i.e., those committed to fostering use) to do evaluation activities in a certain way (e.g., using appropriate and credible methods) in certain contexts (i.e., environments where potential users can take action based on the evaluation process/results). Such a "tendency statement" makes no guarantees, but the commonality of the pattern, supported by research and hinting at a realist framing of "what works for whom in what settings," highlights the general tenets of what the field has learned over roughly five decades of research on evaluation use.

The Centrality of Use in Evaluation Practice

Time has been kind to the concept of evaluation use. Solid evidence now exists detailing its prominent role and widespread acceptance in the field. Whereas historically the field began with a laser focus on methodology, the idea of use is now integral to the way that most practicing evaluators think about evaluation. Evidence of the centrality of use comes from three sources: (1) its prominence in the field's professionalizing documents, (2) its distinct role in numerous evaluation theories, and (3) researchers' continuing focus on this topic.

First, evaluation use has clearly established a visible presence in the evaluation field's professionalizing documents. *The Program Evaluation Standards* (Yarbrough, Shulha, Hopson, & Caruthers, 2010), the newly approved draft of American Evaluation Association (AEA, 2018a) Evaluator Competencies, and the proposed revision of the *AEA's* (2018b) *Guiding Principles* are the oldest and the newest such documents in the field, and, in each, evaluation use/utility plays a role. *The Program Evaluation Standards* (Yarbrough et al., 2010), originally published in 1981 and revised in 1994 and 2011, place utility first of the initial four domains.[1] Utility refers to the extent to which use is possible. Utility " . . . describes when and how evaluation worth is created, for example, when evaluations contribute to stakeholders' learning, inform decisions, improve understanding, lead to improvements, or provide information for accountability judgments" (Yarbrough et al., 2010, p. xxviii).

The eight utility standards reflect two categories from the pattern theory of research-based factors related to increased evaluation use discussed above: (a) evaluator factors, including evaluator credibility, attention to stakeholders (engagement), and concern for consequences and influence and (b) evaluation factors, including meaningful processes (that contain negotiated purposes and explicit values) and products (relevant information) and timely and appropriate communicating and reporting. When the joint committee working on the third edition debated a major reordering of domains that would have put accuracy as the lead domain, Stufflebeam (as cited in Patton, 2012) wrote in an impassioned letter to the joint committee members:

> The new sequencing of the categories of standards is illogical and a counterproductive break with both the JC Standards' historic rationale and the rationale's position in mainstream evaluation thinking and literature . . . The re-sequencing of categories of standards ignores the historic case for the original sequencing of categories of standards, as Utility, Feasibility, Propriety, and Accuracy. Originally, that sequencing was recommended by Lee Cronbach and his Stanford U. colleagues . . . Given the scarcity of resources for evaluation, studies should be conducted only if they will be used. (p. 389)

Following considerable debate, utility remained the first domain.

Table 1. The Use Competencies in the Draft American Evaluation Association Evaluator Competencies.

Domain	Number	Competency
Professional practice	1.12	Fosters the use and influence of evaluation and its results
Methodology	2.15	Articulates evidence-based judgments and recommendations for use
Context	3.10	Attends to evaluation use and influence in context
Planning and management	4.5	Plans for evaluation use and influence
Interpersonal	5.2	Values and fosters constructive interpersonal relations foundational for professional practice and evaluation use

Second, more recently and following 3 years of extensive deliberation, AEA's Competencies Task Force placed a competency related to evaluation use in each domain of the recently board-approved competencies: professional practice, methods, context, planning and management, and interpersonal. Table 1 documents that all five domains of the AEA's competencies highlight the role of evaluation use in competent evaluation practice, signifying its centrality to good practice. The pattern theory's evaluator factor concerning "dedication and commitment to facilitating and stimulating use" aligns strongly with all five of these competency domains.

In addition, other nonuse competencies align with the research-based factors that comprise the pattern theory of evaluation use. The evaluator factor related to developing rapport and a good working relationship with users, for example, aligns with Competencies 5.3 ("Uses appropriate social skills to build trust and enhance interaction for evaluation practice") and 5.7 ("Facilitates constructive and culturally responsive interaction throughout the evaluation"). A second evaluator factor, addressing the way that evaluators involve or engage potential users along with the user factor concerning their meaningful involvement in evaluation processes, aligns with Competency 5.2 ("Values and fosters constructive interpersonal relations foundational for professional practice and evaluation use"). Similarly, the appropriateness of the methods employed and their credibility with potential users aligns with Competency 2.7 ("Designs sound, credible, and feasible studies that address evaluation purposes and questions"), communication quality aligns with two competencies (4.10, "Communicates evaluation processes and results in appropriate, timely, and effective ways," and 5.6, "Communicates in meaningful ways throughout the evaluation . . ."), and concerns about larger organization issues align with three competencies (3.2, "Respects and responds to the uniqueness of the context," 3.3, "Addresses systems and complexity within the context," and 3.9, "Considers both specific and broader contexts of the program"). These direct connections make obvious the centrality of evaluation use considerations in competent evaluation practice detailed in the AEA Evaluator Competencies.

Third, while earlier versions of the *Guiding principles* of the AEA (2018b) made minimal reference to an evaluator's role in fostering evaluation use, one section of the recent revision implicitly attends to issues of use. The section heading of Section E of the proposed revision, "Common Good and Equity," reads: "Evaluators strive to contribute to the common good and advancement of an equitable and just society." Two principles in particular highlight activities related to potential use or—more likely—to avoiding misuse:

- "E3. Identify and make efforts to address the evaluation's potential risks of exacerbating historic disadvantage or inequity
- E5. Mitigate the bias and potential power imbalances that can occur as a result of the evaluation's context." (AEA, 2018b)

Because potential users are often people with power to make decisions that affect others, these revised guiding principles suggest that evaluators need to attend to the use or misuse of evaluation that could lead to negative consequences in terms of social justice.

Taken together, these professional documents provide the first type of evidence of the centrality of evaluation use. The second is provided by the distinct role that use plays in extant evaluation theories. A more detailed discussion of these theories follows below, but unmistakably, the concept of use is an integral part of each of the theories placed on the use branch of the evaluation theory tree (Alkin, 2013); this is, after all, why they were placed on that branch.[2] All of these theories share a common commitment to the evaluator's role in facilitating and stimulating use and actively engaging potential users in the evaluation process. Indeed, Patton (1978) coined the term *personal factor*, highlighting the critical role of intended users in fostering evaluation use; Alkin has written of the *evaluator personal factor*, which includes a strong commitment to the evaluator's role in attaining use (Alkin, Daillak, & White, 1979; Alkin & Vo, 2018), and King and Stevahn (2013) coined the term *interpersonal factor*, emphasizing the need for evaluators to nurture meaningful interactions with evaluation clients and participants en route to subsequent use. Given their practice in the federal policy-making arena, Alkin (2013) notes that Chelimsky and Wholey stress the importance of political sensitivity and credibility and, along with Alkin and Preskill, carefully attend to the characteristics of programs and the nature of the organizations in which evaluations take place.

While a commitment to fostering use is necessarily true for theorists placed on the use branch, concern for the use of evaluation is also integral to theories placed on the remaining two branches. Consider one example from each nonuse branch to make the case: Huey Chen from the methods branch and Donna Mertens from the values branch. Chen's theory, placed on the methods branch owing to its theory-driven orientation, also reflects a sincere commitment to use. The second edition of his textbook, *Practical Program Evaluation* (Chen, 2015), explicitly describes how to work actively with stakeholders on a program's scope and action plan and how to implement a contextually appropriate evaluation explicitly framed to foster the use of the results. His content throughout suggests the importance of interpersonal processes to engage evaluation participants (e.g., participatory evaluation, facilitation, and consensus building). There is a clear commitment to what he calls stakeholder credibility, that is, making the evaluation believable and useful, in addition to scientific credibility, again reflecting a commitment to fostering use. Located on Alkin's values branch,[3] the transformative theory of Mertens (2013), whose "theoretical strands include . . . feminist theories, critical theory, critical race theory, disability rights theory, indigenous rights theory, queer theory, and deafness rights theory" (p. 229), engages the evaluation process to directly target social injustice and inequities. Evaluation use is a necessary component of her approach; without use of both the evaluation process and its outcomes, society will remain unchanged and injustice will continue. Mertens' approach requires that evaluators actively engage members of diverse communities throughout the evaluation process both to understand multiple perspectives and to ensure the appropriateness of methods and their perceived credibility with potential users.

The third source of evidence of the centrality of use comes from the fact that scholars routinely continue to study diverse aspects of evaluation use, including different geographic contexts, programmatic subject areas, and sectors. Further, as will be discussed in more detail below, it is evident that the concept of evaluation use continues to evolve as scholars and published research pay attention to increasing their clients' and stakeholders' focus on what to do both as the evaluation takes place and after it is completed.

Theories of Evaluation Use and Influence Over Time

If there is now little question of the importance of evaluation use in the field, a question does arise as to what extent this centrality is reflected in evaluation theory. The following section begins with a

comment on evaluation theorizing in general, presents a set of criteria for meaningful theories of evaluation use, and applies these to utilization-focused evaluation (UFE), one of the best described prescriptive theories. It then describes numerous theories and models of evaluation use and influence.

A Comment on the General State of Evaluation Theorizing

While the state of evaluation theorizing in general is not the focus of this article, to discuss theories of evaluation use and influence first requires taking a step back to describe the current state of evaluation theorizing as we see it. The point can be made quickly: Whether called a field of practice, a discipline, a transdiscipline, or a profession, program evaluation has yet to develop a singular, overarching, unifying theory or competing theories as is common in traditional social science fields. As Stufflebeam and Coryn (2014) put it, evaluation scholars have not by and large to date engaged in the systematic pursuit of validated social science theory:

> Although evaluation theorists have advanced creative and influential models and approaches for con-
> ducting program evaluations, these constructions have not been accompanied by a substantial amount of
> related empirical research. Consequently, no vast body of evidence exists on the functioning of different
> evaluation approaches ... [T]he program evaluation field lacks a sufficient body of research and steadily
> improving theories flowing from an ongoing process of rigorous, empirically grounded theory develop-
> ment. (p. 46)

King and Stevahn (2013) give several reasons for this, including an alternative focus on program theory, lack of funding, and evaluators' commitment to advancing practice. In the 1960s, as the need for evaluation developed rapidly with little available accompanying research, scholars writing on the topic generated a pragmatic solution: If traditional social science research had not occurred (for whatever reason), better to broaden the definition of theory in evaluation. Lacking research-based theories to guide evaluation practice, the field's intellectual leaders *prescribed* their thoughts on how to conduct evaluation, and these "prescriptive theories" have captured the field's attention.

Evaluation scholars (e.g., Alkin, 2013; Stufflebeam & Coryn, 2014) continue to make a clear distinction between evaluation theory and evaluation models or approaches; however, given the lack of validated theories, the distinction blurs in practice. As Alkin (2013) writes, "... [I]n the strictest sense, what we will refer to as 'evaluation theories' do not fully qualify for that status. Nevertheless, we intentionally refer to them in that way to reflect their most common current usage ... " (p. 4). This usage expands the definition of theory by dividing it into "two general types of models":

> (a) a *prescriptive model*, the most common type, is a set of rules, prescriptions, prohibitions, and guiding
> frameworks that specify what a good or proper evaluation is and how evaluation should be done ... and
> (b) a *descriptive model* is a set of statements and generalizations that describes predicts, or explains
> evaluation activities—such a model is designed to offer an *empirical theory*. (Alkin, 2013, pp. 4–5)

The distinction enables scholars in the field to capture and study the "creative and influential models and approaches for conducting program evaluations" that have emerged over time (Stufflebeam & Coryn, 2014, p. 46). Evaluation "theory" encompasses both traditional social science theory (also known as [aka] descriptive theory) and the distinctive approaches and models derived from practice, often identified with the individual who initially described them (e.g., Stake, Scriven, and Mertens). In the words of Donaldson and Lipsey (2006, p. 59), "... descriptive theories characterize what is and prescriptive theories articulate what should be." We will use this theoretical distinction in discussing theories of use and influence.[4]

Table 2. A Summary of the Use Component.

(1) Addresses all three use bases, including the following: 　　a. A description of possible kinds of use 　　b. A depiction of the time frames in which use occurs 　　c. An explanation of what the evaluator can do to facilitate use under different circumstances (2) Recognizes that 　　a. Context: Use of evaluative results can threaten entrenched interests 　　b. Content: Certain types of information are harder to use than others 　　c. Context: The change implies slow incremental nature of policy that instrumental use is also slow and incremental 　　d. User: Policy-makers often give ideology, interest, and feasibility a higher priority than evaluation results 　　e. User: Using evaluation results is not a high priority for many practitioners who assume the efficacy of what they do 　　f. Evaluator actions: Different activities facilitate different kinds of use, but limited time and resources make it hard to do them all (3) Identifies key choices that evaluators must make in deciding to try to produce useful results 　　a. How 　　b. When 　　c. Where 　　d. Why

Source. Shadish et al. (1991, Table 2.4, p. 53).

Criteria for Meaningful Theories of Evaluation Use—And How They Are Not Applicable

A good evaluator understands the value of explicit criteria in a meaningful evaluation, and, thankfully, authors have proposed two sets of such criteria, one that is applicable for analyzing both prescriptive and descriptive theories, identifying the "critical aspects of the value of theory to practice" (Miller, 2010), and one for a good descriptive evaluation theory overall (Shadish, Cook, & Leviton, 1991). Miller (2010, pp. 391–396) presents five criteria that speak directly to the value of evaluation theory to practice: operational specificity, range of application, feasibility in practice, discernible impact, and reproducibility. Shadish, Cook, and Leviton (1991) propose broader criteria for a "good theory for social program evaluation," including one section specific to evaluation use (see Table 2).

Table 3 matches Miller's criteria with the elements of Shadish, Cook, and Leviton's good evaluation use theory. The match is not entirely one-to-one. For example, Shadish, Cook, and Leviton focus on negative aspects of discernible impact and do not include reproducibility in their list at all, although it is implied by their focus on descriptive theory. Nevertheless, putting the two sets of criteria together and applying them to evaluation use yields a framework for examining examples of use theory over time:

(1) Operational specificity: Explicit details are given about how to foster evaluation use for studies in specific settings.
(2) Range of application: Explicit description is provided of where the theory is likely to increase use and where it is not likely to succeed.
(3) Feasibility in practice: Practitioners can easily and routinely conduct the activities.
(4) Discernible impact: The prescribed activities do, in fact, lead to increased use.
(5) Reproducibility: Different practitioners can reproduce the same outcomes (i.e., use) at different times and places.

Unfortunately, owing to the lack of formal descriptive theories of evaluation use, to apply these criteria is essentially to ensure that the extant "theories" will fail to meet them. Consider UFE, a

Table 3. Criteria for Reviewing Evaluation Use Theories Based on General Criteria From Miller (2010) and Use Theory Criteria From Shadish et al. (1991).

Miller's Criteria (2010)	Miller's Criteria Applied to Evaluation Use	Shadish et al. (1991) Use Theory Criteria
Operational specificity	Explicit details given about how to foster evaluation use (for studies in specific settings)	• A description of possible kinds of use (1a) • An explanation of what the evaluator can do to facilitate use under different circumstances (1c) • Identification of key choices that evaluators must make in deciding to try to produce useful results: How, when, where, and why (3a, 3b, 3c, and 3d)
Range of application	Explicit description of where the use theory is likely to increase use and where it is not likely to succeed	• A depiction of the time frames in which use occurs (1b) • Identification of key choices that evaluators must make in deciding to try to produce useful results: when and where (3b, 3c)
Feasibility in practice	Practitioners can easily and routinely conduct the activities	• Recognition that certain types of information are harder to use than others (relates to users, rather than evaluation practitioners; 2b) • Different activities facilitate different kinds of use, but limited time and resources make it hard to do them all (2f)
Discernible impact	The prescribed activities do, in fact, lead to increased use	Recognition of reasons why use may not occur: • Use of evaluative results can threaten entrenched interests (2a) • The slow incremental nature of policy change implies that instrumental use is also slow and incremental (2c) • Policy-makers often give ideology, interest, and feasibility a higher priority than evaluation results (2d) • Using evaluation results is not a high priority for many practitioners who assume the efficacy of what they do (2e)
Reproducibility	Different practitioners can reproduce the same outcomes (i.e., use) at different times and places	—

prescriptive theory that is the subject of five lengthy books—four editions of *Utilization-focused Evaluation* (Patton, 1978, 1986, 1997, 2008) and a 440-page primer, *Essentials of Utilization-focused Evaluation* (Patton, 2012). Surely, given the amount of content available on UFE, if any theory of use could meet the criteria proving its value to practice, it might be this one. Analyzing the 17 UFE steps listed in *The Essentials of UFE*, it does fairly well on four of the criteria:[5]

- Operational specificity: Patton (2012) identifies the 17 steps of UFE, providing explicit details, step-by-step, about how to plan and implement evaluations that are likely to foster

use in specific settings. He makes it clear that, given the context-oriented nature of any evaluation, there can be no guarantees.

- Feasibility in practice: Results of a survey of U.S. members of the AEA (Fleischer & Christie, 2009) document the extent to which practitioners reportedly attend to issues of use, suggesting that evaluators are able to routinely engage in utilization-focused activities (e.g., planning for use at the beginning of a study, identifying and prioritizing intended uses and intended users, involving stakeholders in the evaluation process).
- Reproducibility: Novices learning about UFE often lament the fact that they lack the intellectual and interpersonal skills of Michael Quinn Patton. "I'll never be as good as he is," they moan. "His skill set can't be reproduced!" But what exactly would it mean for different UFE practitioners to produce the same outcomes (i.e., use) at different times and places? It seems possible to us that different UFE proponents could effectively engage in the practice at different times and in different contexts and successfully foster use. The challenge for this criterion may well rest in what we mean by "use" and how we measure it.
- Range of application: In his UFE writing, Patton clarifies settings where use is likely and where it is not likely to succeed. The first step of UFE—"assess and build program and organizational readiness for utilization-focused evaluation" (Patton, 2012, p. 15)—addresses this directly. If a program or its organization is not ready for evaluation, the implication is that use is unlikely to occur. Step 3—"identify, organize, and engage primary intended users" (PIUs; Patton, 2012, p. 61)—extends this to the research-based notion of the personal factor, that is, the importance of engaging people who actually are interested in and care about the evaluation process and its results. We should note, however, that Patton's theory does not make a distinction among types of evaluation—it applies to all—nor does it make a distinction as to whether the theory is equally valid for large-scale versus small-scale programs, calling into question its range of applicability.

If UFE mostly meets four criteria for a useful theory of use, however, it certainly falls short on the remaining criterion: discernible impact. To say with confidence that prescribed UFE activities *consistently* (i.e., *always*) lead to increased use is to ignore the contextual nature of evaluation settings where, for example, the elimination of a PIU, UFE's Achilles' heel (Patton, 2008, pp. 566–567), can scuttle even the most masterful UFE study. As noted previously, there simply are no guarantees. And if this most detailed of prescriptive theories cannot meet all five criteria for a meaningful theory of evaluation use, then other less detailed theories will surely fail to meet them.

Having a set of criteria for evaluating use theories, however, is surely helpful even it reaffirms the fact that the field is a long way from having an empirical-grounded descriptive theory of evaluation use. On the one hand, at least scholars know what theory is needed ultimately. On the other hand, however, absent such a descriptive theory (or theories), evaluation practitioners may continue to adopt and adapt prescriptive theories with which they are familiar without knowing exactly what to do to increase the likelihood of potential use. In the meantime, given the absence of a descriptive use theory and perhaps to stimulate its eventual creation, it makes sense to briefly trace the evolution of prescriptive use theories to identify their distinctive attributes.

Evaluation Use Theories

In the spirit of historical documentation, rather than analyzing use theories by applying stringent criteria, we will instead first describe their evolution, explicate what distinguishes them, and finally examine two formal theories of use.

Prescriptive use theories. Alkin (2013) places nine theorists on the use branch in the following order: Daniel Stufflebeam, Joseph Wholey, Eleanor Chelimsky, Marvin Alkin, Michael Quinn Patton,

David Fetterman, Hallie Preskill, Jean King, and J. Bradley Cousins. As one of the first theorists to frame the evaluation process to foster use, Stufflebeam is placed on the branch itself near the base of where it branches off from the trunk, with Patton, whose writings since the late 1970s have influenced generations of evaluators, farther out on the branch. Both are featured as checklists on the Western Michigan University's Evaluation Center website (https://wmich.edu/evaluation/check lists). The remaining seven theorists each occupy a leaf on the use branch. Table 4 provides a brief summary of these theorists' writings over time, highlighting why they belong on the use branch. What features do they share? They all make evaluation use central to the design and implementation of evaluations, paying careful attention to context, situation dynamics, and engaging potential users in hopes of fostering various kinds of use, both of its process and results.

Formal models of evaluation use. In addition to the nine prescriptive theories described in Table 4, scholars have proposed two theoretical models that explicitly address evaluation use:[6] (a) a theoretical model of evaluation utilization (Johnson, 1998) and (b) the ecological model of evaluation use (Ottoson & Martinez, 2010).

First, about 20 years ago, Johnson (1998) reviewed both what he called implicit and explicit "evaluation utilization process models" to propose a unifying theoretical model of evaluation utilization, much like an outcomes chain for the use process. "In an implicit process-model, variable ordering or process is implied but is not directly depicted by the evaluator" (Johnson, 1998, p. 95). To put these models together, Johnson had to make multiple assumptions and add his own thinking since the work is based on the authors' writing but not confirmed. In this way, he created implicit models for seven theorists: Donald Campbell, Michael Scriven, Carol Weiss (two versions), Joseph Wholey, Robert Stake, Lee J. Cronbach, and Peter Rossi. Next, he presented nine additional process models, that is, explicit models that were "constructed by researchers and generally, appear in articles and books" (p. 97), to highlight the theoretical contributions of 10 authors, including Jennifer Greene, Marvin Alkin, and Michael Quinn Patton.[7]

Johnson then took the 16 models he created, added the concept of organizational learning, which did not appear in any of them, and compiled a single unified theoretical model of evaluation utilization "intended to apply to any evaluation" (p. 106), surely a serious claim (Figure 1). The model has multiple components, presented as variables suitable for research:

- The external and internal environments/contexts of the evaluation, which give inputs to and receive outputs from the chain of logic;
- background variables, including organizational, individual, and evaluator characteristics;
- interactional variables, comprising evaluation participation, dissemination, and politics;
- explicit utilization variables, divided into three categories, that is, "a multidimensional conceptualization of the outcome variable" (p. 103), including cognitive use and behavioral use and adding organizational learning, a variable not present in any of the process models, but mentioned in the literature.
- Johnson's model suggests that "... evaluation use occurs through an open system of interrelated background, interaction and use variables operating in an internal environment situated in an external environment..." (p. 106). He emphasizes that it is not a static model. Rather, it is a "'model in action'... based on the assumption that the utilization process needs to be viewed as a dynamic and open, complex system..." (p. 107). Table 5 compares Johnson's variables with the pattern theory of evaluation use.

Whereas Johnson (1998) created his model by integrating 16 self-created "process models" of evaluation utilization, Ottoson and Martinez (2010) based the second model, their ecological model, on the results of a single case study of use, a limited source of data for a generic model. They

Table 4. Evaluation Use Theories From the Evaluation Theory Tree.

Theorist	Approach Name or Associated Term	Decade of Origin	Reason for Placement on the Use Branch	Distinction Features of the Approach
Daniel Stufflebeam	Context, input, process, and product	1960s	Focuses on people's use of evaluation for decision-making and accountability	• It is framed around four concepts—context, input, process, and product—that apply to all evaluation settings • "...[E]valuation's purpose is not only to prove but also to improve" (Stufflebeam, 2013, p. 244)
Marvin Alkin	Context-sensitive evaluation	1960s–1970s	Grounds the evaluation process in context, including evaluators' commitment to use and engaging and training primary users/stakeholders to foster use of the process and results	• Evaluators must systematically seek context-specific information (individual, organizational, community, cultural) and use it to inform the evaluation process • The "*valuing process must be user engaged*" (Alkin & Vo, 2018, p. 300, emphasis in original) • Evaluators identify, engage, and train users/stakeholders who have the interest and possibility to use the results
Eleanor Chelimsky	—	1960s–1970s	Pays attention to an evaluation's purpose and situation dynamics to create credible information for decision makers in highly political settings	• Evaluators must work actively to remain credible in politically charged settings • The three purposes for evaluation—(1) accountability, (2) knowledge acquisition, and (3) management/development—affect the evaluator's role in specific situations • "...asking skeptical questions about conventional wisdom" can be important (Chelimsky, 2013, p. 278); it is important to speak truth to power
Joseph Wholey	Sequential purchase of information	1970s	Designs studies to provide decision makers useful information from beginning to end	• This approach provides "...retrospective assessment of the performance of programs...that have been implemented by public and nongovernmental organizations" (Wholey, 2013, p. 261) • The approach includes evaluability assessment, rapid-feedback evaluation, performance measurement systems, and impact evaluation
Michael Quinn Patton	Utilization-focused evaluation	1970s	Frames every evaluation around use by engaging primary intended users from start to finish	• "Evaluation done for and with specific primary intended users for specific, intended uses" (Patton, 2008, p. 37) • Evaluations should be "judged by their utility and actual use" (Patton, 2012, p. 4) • The approach centers on situational responsiveness and an evaluator who is "active, reactive, interactive, and adaptive" (Patton, 2012, p. 197)

Table 4. (continued)

| Theorist	Approach Name or Associated Term	Decade of Origin	Reason for Placement on the Use Branch	Distinction Features of the Approach
Hallie Preskill	Evaluative inquiry for learning in organizations	1990s	Believes evaluation should be conducted only if individuals, groups, and organizations will learn from it	• Learning at multiple levels is the key purpose of evaluation • It is important to actively engage people in the evaluation process and build their capacity to engage in evaluative thinking over time • Evaluation should be embedded in organizational structures and functions
Jean King (with Laurie Stevahn)	Interactive evaluation practice	1980s–1990s	Purposefully engages people in interactions that promote the evaluation process and its use	• Evaluators should work to foster evaluation use by engaging people in the process • Relationships are critical for both evaluation use and evaluation capacity building over time • Participation in evaluation should be a learning experience for people • Sustaining evaluation within organizations is an evaluator's ultimate and extremely difficult goal
J. Bradley Cousins	Participatory evaluation	1980s–1990s	Engages stakeholders as a way to privilege use and overcome the problem of nonuse	• There is a distinction between practical participatory evaluation, which works on local change, and transformative participatory evaluation, which attends to issues of social justice • It is critical to integrate evaluation into organizational functions • His research has focused on understanding participatory evaluation and how to implement it effectively in organizations
David Fetterman	Empowerment evaluation	1990s	Engages multiple participants in activities that integrate use into the evaluation process	• Participants learn how to think like an evaluator, that is, the process teaches people evaluation skills • Three-step approach: (1) establish mission, (2) take stock of current status, and (3) plan for the future • Its goal is to give people the power to develop and advance their own programs
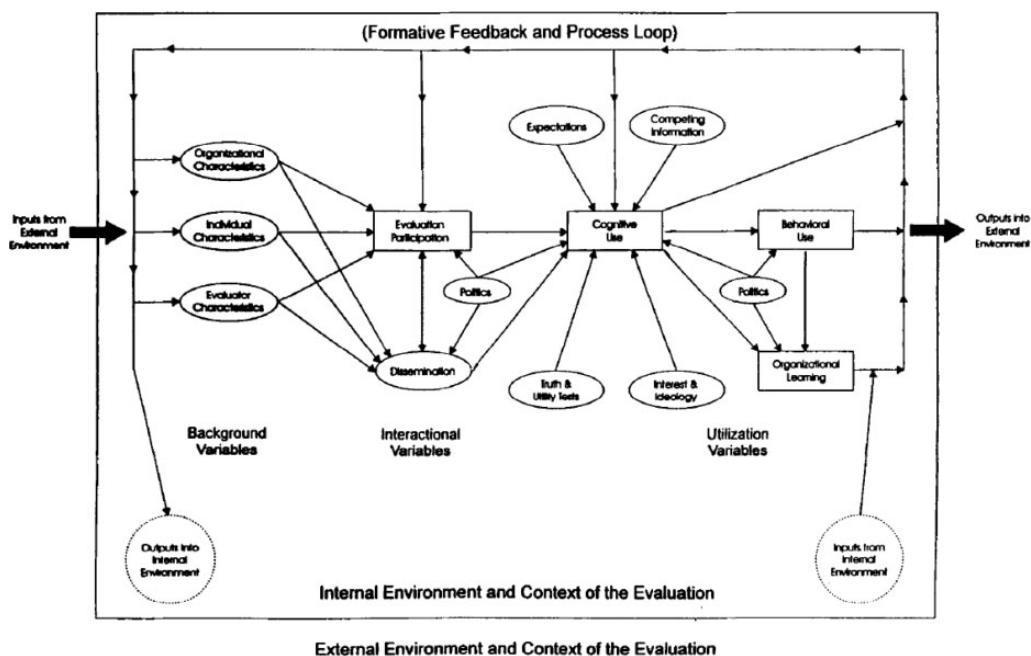
Figure 1. A theoretical model of evaluation utilization (Figure 19 in Johnson, 1998, p. 104).

Table 5. A Comparison of the Pattern Theory of Use and Johnson's Theoretical Model of Evaluation Utilization.

Pattern Theory Categories	Johnson's Model Content
User factors	Background: individual characteristics
Evaluator factors	Background: evaluator characteristics
Evaluation factors	Interactional variables:
	• Evaluation participation
	• Politics
	• Dissemination
Organizational/social context factors	Background: organizational characteristics
	Internal environment and context of the evaluation
	External environment and context of the evaluation

interviewed a total of 23 informants—"staff, evaluators, grantees, program developers, consultants and other stakeholders" (p. 6)—from the 4-year evaluation of an R. W. Johnson Foundation-funded program entitled Active for Life: Increasing Physical Activity Levels in Adults Age 50 and Older. As they put it,

> Linear models just do not tell the interactive story we found. Like other ecological models, this one proposes multiple "eco-systems" or contexts of evaluation use in this case study, with multi-directional and multi-layered influences. (p. 7)

The resulting ecological model (see Figure 2) consists of four concentric circles surrounding the bull's-eye, which is labeled "valuing," the "core work of evaluation" (p. 7). In order, from inside to out, the circles are labeled immediate program, community, field, and society. Ottoson and Martinez
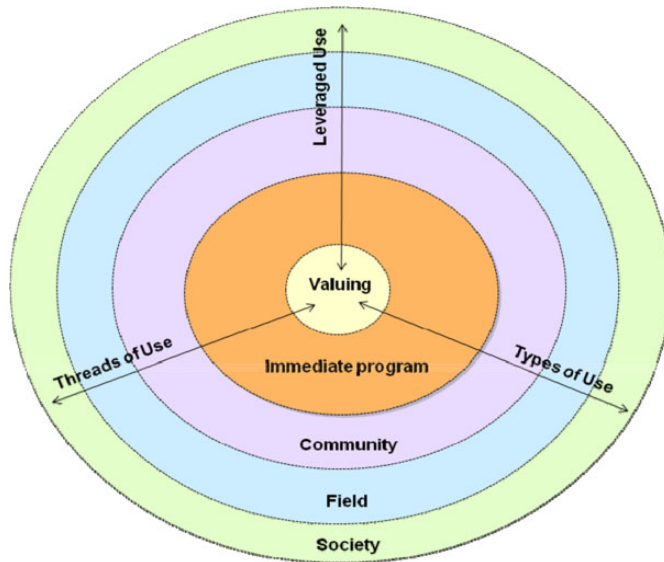
Figure 2. The ecological model of evaluation use (Ottoson & Martinez, 2010, p. 8).

(2010) present the case study results as "multiple paths to understanding evaluation use," with three components that cut across the four circles to the valuing core of the model: types of evaluation use; and "threads of use" and "leveraged use," two components that "more fully tell the story of use" (p. 15) and smack of Kirkhart's concept of influence.[8] Given its emphasis on use, this model emphasizes the organizational contexts affecting evaluation use but does not highlight the pattern theory's user, evaluator, or evaluation factors.

Taken together, the two models present different images of the use process. They minimally fall into the descriptive theory category: One builds on researcher-created models based on other theorists' writings; the other is based on case study data from a single study. Grounded in internal and external environments/contexts, Johnson's linear model, with the obvious limitations of such thinking given the complexities of social programs, moves purposefully from background variables to interactional variables then to utilization variables. Eschewing a linear approach and, again, grounded in data from a single case study, Ottoson and Martinez, by contrast, nest evaluation in its layers of surroundings, moving from the center outward.

But while taking different approaches, these models share commonalities. Each points to multiple and complex possibilities for use and to the systemic and multivaried nature of the use process, including the role of politics and interpersonal dynamics. Each also makes it clear that evaluation use can occur in different ways and over time. Either potentially offers a framework for systematic research on the topic.

Theories of Evaluation Influence. At the turn of the 21st century, Kirkhart (2000) proposed a new term—*evaluation influence*—as "the capacity or power of persons or things to produce effects on others by *intangible* or *indirect* means" (p. 7, emphasis added). The discussion of such a shift in the early 2000s led to an upsurge of discussion of the consequences of evaluation. No historical summary of the evaluation use literature, therefore, would be complete without discussing the two published "theories" of influence: (a) the integrated theory of evaluation influence (Kirkhart, 2000) and (b) the schematic theory of evaluation influence (Henry & Mark, 2003; Mark & Henry, 2004). It

is interesting to note that, in contrast to the models of evaluation use that were labeled *models*, these authors labeled their work *theories*. If, as Bacharach (1989) writes, "The goal of theory is to diminish the complexity of the empirical world on the basis of explanation and prediction" (p. 513), these are not traditional social science theories; they are not grounded in empirical research. Calling them theories instead applies the field's tradition of equating "theory" with models or approaches (Alkin, 2013).

Integrated theory of evaluation influence. Kirkhart (2000) argued that to understand how evaluation actually changes society, researchers should extend the narrow framing of use by adding a broader-based construct. She suggested the term *influence* as an addition to *use*:

> The term *influence* . . . is broader than use, creating a framework with which to examine effects that are multidirectional, incremental, unintentional, and noninstrumental, alongside those that are unidirectional, episodic, intended, and instrumental. (which are well represented by the term *use*; Kirkhart, 2000, p. 7, emphasis in original)

The more expansive term, she reasoned, would foster a more "inclusive understanding of the impact of evaluations" (Kirkhart, 2000, p. 5). Kirkhart's theoretical model has three dimensions: intention (intended or unintended), source (evaluation process or results), and time (immediate, end-of-cycle, and long-term). The 12 resulting possibilities clearly include forms of evaluation use, for example, intended outcomes that develop when the evaluation results become available at the end of an intervention or unintended outcomes that occur immediately as a result of an evaluation process (aka process use). But Kirkhart (2000) feared that researchers' limited focus on use might exclude other possibilities (e.g., unintended outcomes resulting from a long-term evaluation process or the long-term effects of specific evaluation findings in a policy setting). "Taken together," she wrote, "the three dimensions of influence offer a framework within which to examine both the positive and negative impacts of evaluation" (p. 13).

Alkin and Taut (2003) expanded Kirkhart's model by adding the element of *awareness*, purposefully focusing attention on what evaluators can be aware of and do something about, that is, intended and unintended use that is immediate and end-of-cycle. They called these instances evaluation use. Any impacts that are unintended and of which the evaluator is unaware, they wrote, appear "not as essential to the evaluation profession as the impacts that are of a conscious . . . nature, in the eyes of the users, and hopefully of the evaluator as well" (p. 9). They referred to the instances that are long-term and beyond the control and the purview of the evaluator as influence.

Schematic theory of evaluation influence. Building on Kirkhart's argument, Henry and Mark published "Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions" in 2003 and, with the author order reversed, "The Mechanisms and Outcomes of Evaluation Influence" the following year. In the first article, they joined Kirkhart in moving beyond the concept of use, writing that " . . . neither the change processes through which evaluation affects attitudes, beliefs, and actions, nor the interim outcomes that lie between the evaluation and its ultimate goal—social betterment[9]— have been sufficiently developed" (Henry & Mark, 2003, p. 293). They sought to identify these change processes and outcomes by expanding the scope of the field's scholarly content:

> Fortunately, the framework need not be developed de novo. Social science provides both theories of change that are relevant and research on specific outcomes that are similar to those that can be expected to appear in chains of outcomes through which evaluation could lead to social betterment. (Henry & Mark, 2003, p. 296)
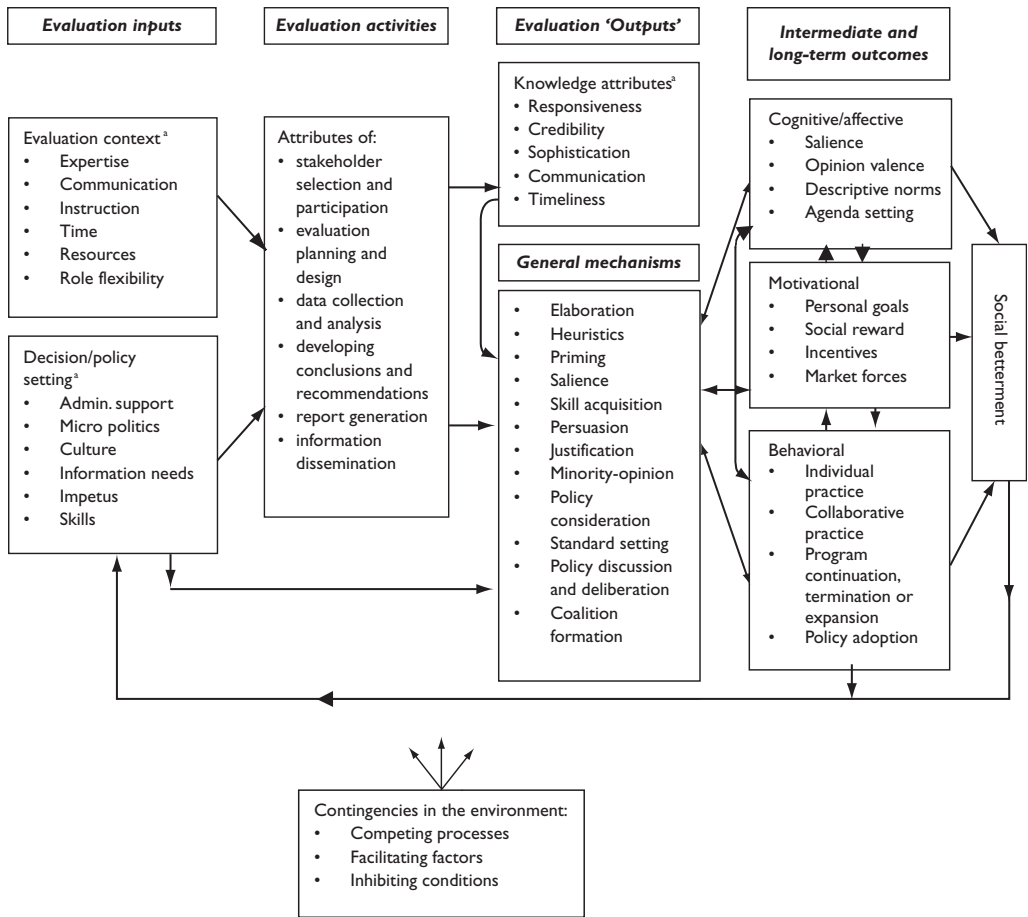
Figure 3. The schematic theory of evaluation influence (Mark & Henry, 2004, p. 46).
[a]Selected elements from Cousins (2003). Categories in bold are taken from Table 1 in Mark & Henry (2004); see table for complete listing.

Although Henry and Mark did not define influence explicitly (Nunneley, 2010; Nunneley, King, Johnson, & Pejsa, 2015), they identified 15 specific mechanisms selected from existing social science research literature (e.g., skill acquisition, social norms, policy change) through which evaluation can produce influences via multiple pathways at three levels: individual, interpersonal, and collective. They concluded by noting that "[b]y combining change processes into causal pathways, we have assembled a more powerful way to formulate working hypotheses about evaluation outcomes" (Henry & Mark, 2003, p. 311).

In their second article, Mark and Henry further developed their ideas by presenting a model that crosses four types of processes/outcomes (general influence, cognitive and affective, motivational, and behavioral) with three levels of analysis (individual, interpersonal, and collective) to categorize selected alternative influence mechanisms, many of which were included in the first article. Building on Cousins (2003), they presented a refined graphic of a theory of evaluation influence that integrated their ideas (see Figure 3). It is a functional logic model for program evaluation and includes inputs, activities, outputs, general mechanisms, intermediate and long-term outcomes in three areas (cognitive and affective, motivational, and behavioral), and contingencies in the environment, all

culminating in social betterment. Again, with no explicit definition of influence, they argued that this "schematic theory" would enable researchers to identify and study specific influence pathways and make "concrete predictions about the general relations between different components of the logic model of evaluation" (Mark & Henry, 2014, p. 47).

At the time these influence theories were presented, they seemed to provoke the field to think more broadly about the potential bandwidth of evaluation use. Kirkhart's theory sought to identify the long-term effects of program evaluation and their sources; Henry and Mark's theory, grounded in realist thinking, sought to identify mechanisms from a variety of disciplines (e.g., behavioral and social psychology, management science, and political science) that could identify pathways to explain the ultimate impact of evaluation. Miller's criteria for a good evaluation theory (i.e., operational specificity, range of application, feasibility in practice, discernible impact, and reproducibility) make clear how influence genuinely differs from use. By determining the causal pathways that lead to observable effects, it surely focuses on impact but little on explicit evaluation practice that might lead to it. As Alkin and Taut (2003) noted, because influence is an *intangible* and *indirect* process, there is no way to specify details of what evaluators should do to increase it.

Influence theory, finally, is more about studying evaluation than about changing its practice. It may help make sense of how evaluations effect change over time, but only when there is a significant body of research on influence might the implications for evaluation practice (i.e., what to do in specific settings to increase the power of evaluation) become evident. That research base on evaluation influence is not yet available, in part because the term *influence* means different things to different people. As Herbert (2014, pp. 388–389) writes, "Despite the prominence of evaluation influence in the literature, there is slow progress toward a persuasive body of literature."

To review, this discussion of evaluation use/influence theory has consisted of three parts. The first reviewed 14 prescriptive theories, each of which makes evaluation use a critical attribute. The second presented two published theories of evaluation use. The third presented two extant theories of evaluation influence. What, now, is the current status of evaluation use theory? In one sense, even after 50 years, evaluation use theory is in its infancy. Speaking about evaluation theory in general, Alkin and Vo (2018) describe the situation succinctly: "There simply is not a sufficient body of knowledge about what happens in an evaluation to be able to predict with certainty what would happen when an evaluation is employed in a particular way" (p. 297). Evaluation use theory is encompassed in that assessment. But in another way, while there is surely a long way to go before scholars develop a meaningful descriptive use theory, we believe that the various "theories" presented here—based on practical experience and to some extent upon research—are helpful beginnings because they focus attention on what we know to be critical attributes related to evaluation use.

Scanning Research on Evaluation Use: Past and Present

Regardless of the state of use theorizing, the fact that researchers have studied this topic since the 1970s is an indicator of its centrality to evaluation practice. The second article in this series (Alkin & King, 2017) provided a history of this research, identifying factors shown to affect evaluation use. In this section, we will discuss a critique of this past research, then—given that this is a historical piece—briefly discuss present and future research, But first a comment on competing understandings of the term *use*.

In contrast to the Inuit people who require multiple words to describe different types of snow, our field has stuck with one term—use—to cover multiple possibilities. Almost 30 years ago, in *Debates on Evaluation*, Alkin (1990) discussed a disagreement, long since resolved, between Weiss (1988a, 1988b) and Patton (1988), noting that a key part of their tension resulted from a difference in focus. Weiss worked in complex policy environments where evaluation results were one piece of information available to decision makers. She thought there had been only "indifferent success in making

evaluation the basis for discussions" (Alkin, 1990, p. 225); decisions accreted, and, if evaluation did not lead directly to decisions, it often led instead to decision makers' enlightenment.

By contrast, Patton worked with PIUs in settings where decision makers with a sense of intended uses might directly apply evaluation results, and he could easily provide many examples of the instrumental use of evaluations. The distinction at that time may have been between so-called academic evaluators and client-centered evaluators, but the crux of the disagreement stemmed from different understandings of what use meant and how someone would know it when they saw it. The ultimate resolution was coming to understand that "evaluation use is not a question of *either* enlightenment *or* instrumental use, but rather *both/and*" (King & Stevahn, 2013, p. 54). Such rival understandings of evaluation use have complicated research efforts, as has the relatively recent addition of the concept of influence. The definition for evaluation use proposed by Alkin and King (2017) sought to create consensus on a definition, one step toward clarifying a path forward with a common understanding. It remains to be seen whether such a stipulation is useful.

Past Research

As noted, Alkin and King (2017) reviewed research on evaluation use factors, including three major compilations (Cousins & Leithwood, 1986; Johnson et al., 2009; Shulha & Cousins, 1997). Although researchers have conducted multiple studies on the topic over the years, Brandon and Singh (2009) raised questions about the quality of the work:

> As a body of evidence for a scientific understanding of the use of evaluation findings . . . the results of the studies on use are currently of questionable quality . . . Standing alone as a body of results about evaluation use . . . the findings of the studies examined here do not as a whole have sufficient scientific credibility. (p. 135)

They based their analysis on 52 studies included in five reviews of research on program evaluation use: two of the three cited above (Cousins & Leithwood, 1986; Shulha & Cousins, 1997), two earlier reviews (Leviton & Hughes, 1981; Thompson & King, 1981), and an article on the "utilization effects of participatory evaluation" (Cousins, 2003).

Their critique applied two criteria: (1) "the *balance of the types of methods* and the implications of this balance for making conclusions about use" and (2) *content-related validity* (Brandon & Singh, 2009, p. 125, emphasis in original). The 52 studies used four methods—"surveys, quasi-experimental simulations, case studies, and narrative reflections" (Brandon & Singh, 2009, p. 133)—and, because 69% were grounded in education, only that area (i.e., the use of education evaluations) met the first criterion. The second criterion "addresses the quality of the methods—that is, the extent to which they showed evidence of content-related validity" (Brandon & Singh, 2009, p. 133), and, taken as a whole, the 52 studies failed: "We found little discussion of content validity issues in the quantitative studies or parallel information in the qualitative studies" (Brandon & Singh, 2009, p. 133).

Nevertheless, Brandon and Singh (2009) conclude that thoughtful evaluators, aware of the limitations, might well apply the research results in ongoing efforts to improve evaluation use in their practice. "This approach implies that utility trumps accuracy . . ." (Brandon & Singh, 2009, p. 134).[10]

Current Research Activity

Critique notwithstanding, people continue to study evaluation use, and a search for articles on evaluation use research will generate an ever expanding list. Table 6 gives examples of studies that scholars have conducted since the Johnson et al. (2009) research compilation, grouped by place, organizational setting, content, and methods. The list is not meant to be inclusive, but illustrative,

Table 6. Examples of Evaluation Use Research Studies (2009–Present).

Category	Examples	Citations
Place	- Brazil - Europe - Ghana - Italy - Israel - New Zealand - Switzerland	- Cornachione, Trombetta, and Nova (2010) - De Laat and Williams (2014) - Akanbang, Darko, and Atengdem (2013) - Rebora and Turri (2011) - Neuman et al. (2013) - Blewden (2010) - Balthasar (2009)
Organizational setting	- Foundations - Government - Internal evaluation - Nongovernmental organizations - School districts	- Thompson and Patrizi (2010) - Cousins, Goh, Elliott, Aubry, and Gilbert (2014); Picciotto (2016) - Loud and Mayne (2013) - Karan (2009); Liket, Rey-Garcia, and Maas (2014) - Sturges (2015)
Content	- Agricultural education - Conservation - Education - Energy - Health promotion - International development - Policy-making - Social psychology	- Lamm, Israel, and Harder (2011) - Jacobson, Carter, Hockings, and Kelman (2011) - Burr (2009) - Lehtonen (2013) - Hartz, Denis, Moreira, and Matida (2009); Yeary, Klos, and Linnan (2012) - Tennant (2010) - Contandriopoulous and Brouselle (2012); Gudmundsson and Sørensen (2013); Højlund (2014) - Fleming (2011)
Methods	- Citation analysis - Q methodology	- Greenseid and Lawrenz (2011) - Baptiste (2010)

and the evidence is clear. Researchers in countries around the world persist in studying questions about evaluation use,[11] and they do so in organizational settings that range from large to small and in programs from a number of different subject areas. Collectively, authors continue to use tried and true methods (e.g., reflective case narratives and surveys) and have added at least two new methods in their studies.

Because the present article is not a formal review of current literature, consider just four articles published since the Johnson et al. (2009) compilation that highlight the range of issues scholars are tackling. They represent research from Canada, Denmark, Israel, and the United States. Contandriopoulos and Brouselle (2012) apply knowledge exchange concepts—polarization and cost sharing—to contrast evaluation models (utilization-focused, realist, empowerment, and democratic evaluation) and to highlight the relationship among context, choice of model, and use of results. In his study, Højlund (2014) uses institutional theory in a policy environment, concluding that "... we need to focus more on the organizational context of evaluation and less on the evaluation and its immediate conditioning factors" (p. 38). Examining the use process in a single educational organization, by contrast, Neuman, Shahor, Shina, Sarid, and Saar (2013) develop a "local theory" about evaluation use specific to that organization and discuss how such theories might help increase use. Sturges (2015) takes a critical stance to present a case study that documents "evaluation's complicity in helping to maintain power asymmetries" (p. 462), offering suggestions to alleviate or at least address such problems.

Where Do We Go From Here?

At the beginning of this article (the final of three), we explained that our purpose in compiling a history of evaluation use was to document and clarify its evolution and, in so doing, suggest the potential of certain practices and contextual considerations for scholars and practitioners. We begin this conclusion by summarizing four ideas that emerged from the review:

1. Evaluation use research has a 50-year history; the concept of evaluation influence has existed for roughly 20 years but has not yet generated a sizable research literature.
2. The research-based pattern theory of use proposed in our second article identified four sets of key factors to study, those related to users, evaluators, the evaluation process, and its context.
3. Scholars continue to conduct research on evaluation use around the world using different framing theories and methods with different types of people in a variety of different settings.
4. At this time, the question of what explicit theory might help evaluators improve practice and increase the appropriate use of the evaluation process and its results remains just that—a question.

Writing about good evaluation theory almost 30 years ago, Shadish et al. (1991) proposed that detailing an evaluation practice component would be most important because " . . . evaluators have to practice in a context where leisurely reflection about theoretical alternatives must yield to action within constraints" (p. 37). Ideally, such a practice theory would identify contingencies, including if/then statements, that is, if an evaluator is in this situation in this context, then here is the best action to achieve the outcome of evaluation use, as defined now in Alkin and King (2017). Based on the initial 50 years of research, it is likely that the development and validation of such a contingency theory, although extremely desirable, may well be the holy grail of use research, especially within the limits of available funding.

How, then, might the field advance thinking on use and influence? Let us begin with two ideas to lay groundwork for a path forward. First, in considering context both small and large, it is important to note how radically times have changed since the 1970s when scholars published the first research on evaluation utilization. It is an understatement to note that the world has changed significantly in the roughly 50 years since people began studying evaluation use, and future research needs to attend to the ways the scene has changed and the impact these changes have had on acts of use and influence themselves and also on how to study these concepts. Consider these points of context:

- Evaluation is a growth industry internationally. The continuing expansion of voluntary organizations of professional evaluators across the globe documents the development and expansion of evaluation activities. There is a wider acceptance of evaluation and expectations that programs will be evaluated and that something should happen as a result. The growing numbers of decision makers with access to evaluation processes and results in organizations ranging from small nonprofits to large multinational corporations increase the possibilities of use, misuse, and influence as social change efforts become increasingly interconnected and complex.
- While evaluation remains an important means of accountability, it has increasingly taken on an additional role as a learning activity. Organizational learning, mainstreaming evaluation, and evaluation capacity building (ECB) reflect the possibilities of "do it yourself" evaluation that is integrated into ongoing program functions with or without support from a professional evaluator. Program staff and administrators can learn evaluation processes and create systems for data collection, analysis, interpretation, and use over time. New roles for evaluators, for example, include teaching workshops on data collection and interpretation, facilitating data

parties, and developing innovative report formats that engage potential users in making sense of evaluation results.

- In these 50 years and reflected clearly in the recent draft revision of *AEA's Guiding Principles*, the field has acknowledged the critical importance of diversity and inclusion and of evaluators' need to focus on issues of power and equity. This is directly related to the broader acceptance of multiple approaches to knowledge creation, including those of indigenous peoples, with Northern Enlightenment approaches contrasting dramatically with those of the global South. The issue of "whose truth for use by whom" and the means for generating it can raise difficult questions for evaluators, especially when they do not reflect the cultural background of program participants. What exactly does it mean for evaluation "to speak truth to power"?

- Perhaps the most visible change in the past 50 years, however, has been in the development of technology that significantly affects how evaluators work. Personal computers, laptops, the Internet, cell phones and their associated apps, Google and other search engines, and almost daily innovations have radically changed how people—evaluators and their clients alike—create and access information. Examples are numerous. Evaluators and potential users now have easy online accessibility to evaluation (and research) studies in different forms. Evaluation reports are available on agency and organization websites rather than in the gray literature that required potential users formerly to write to ask for hard copies. If decision makers around the world have access to technology and the electricity to run it, they have 24/7 access to evaluation information, including reports from multiple funding sources, and they can research how to conduct evaluations, compile studies on similar projects, and even identify evaluators to hire. Evaluators and program leaders can reuse the standardized tools developed for one evaluation that are available online. There are evaluation blogs and webinars and websites from professional evaluation associations. There are networks of users on the evaluation of specific topics (e.g., science education, disaster relief, and environmental change).

- Finally, of grave concern for the practice of evaluation—and an area where we may well feel helpless about what to do individually and collectively—is the recent emergence of an era of "fake news," where experts are distrusted, the scientific method is no longer necessarily valued, and some political leaders reject and even disparage the use of meaningful data. We are reminded anew that the Latin root for the word fact is "to make" (cf. "factory") as people create facts to support a position or fit their desired outcomes.

These contextual changes—the rapid expansion of evaluation, its newfound potential for learning, explicit attention to issues of diversity and equity, technological developments, and the emergence of fake news—occurred over the many years during which evaluation use research took place. We believe it is important to acknowledge current environmental conditions in hopes that scholars will take them more explicitly into account as they plan future studies.

Second, if it is worth reflecting on the effects of an evolving context, it is equally valuable to step back and look at the overarching framing of research on evaluation use. One framework for doing so is to apply Gowin's Vee heuristic (Novak & Gowin, 1984). Gowin developed the Vee to "illustrate the conceptual and methodological elements that interact in the process of knowledge construction..." (Novak & Gowin, 1984, p. 3). At the point of the Vee are the events or objects of interest to the research, in this case, instances of evaluation use or influence. The two legs of the Vee document the methods and the concepts that will guide the inquiry. Ideally, the methods for creating records (data) will clearly match the concepts, so that, taken together, they succeed in answering the research's focus question (which in Gowin's system is written in the center of the Vee).

Applied in brief given the limitations of space, each of the three parts of the Vee furthers this discussion:

- Events of interest: Given the many changes in context just described, it is highly likely that the events of interest around evaluation use/influence have changed since the 1970s, and researchers should attend to this. Below, we will discuss how Alkin's concept of context-sensitive evaluation may be a viable way to focus future research (Alkin & Vo, 2018).
- Methods: As discussed in the section on continuing research, scholars are employing new methods for the study of use/influence, and this methodological innovation should continue as it could well prove useful.
- Concepts: Critics may blame the construct of evaluation use for less than definitive findings after 50 years. Even with the more recent addition of the concept of influence, the lack of definitive answers to questions of use may suggest a causal relation between a construct that is ill-defined and diffuse and the lack of conclusive research results. We would argue, however, that it is not a question of the construct being too broad. Rather, we believe that the problem may stem from the fact that people are missing the distinction between what the literature says is evaluation use and what is evaluation influence, including a lack of specificity of these concepts' outcomes. What exactly does evaluation use or evaluation influence, a type of use, look like in different settings? Given a specific type of organizational setting, how would you know either if you observed it? This was the reason we defined evaluation use in Alkin and King (2017). If scholars can agree on a definition—ours or someone else's—and explicit outcomes of evaluation use and create validated instruments for measuring it across various locales, studies would have common metrics for comparison across contexts.

To summarize, we believe that future research should pay close attention to the evolving contexts of evaluation use and of the need for a common definition and outcomes.

We also want to suggest a family of evaluation use theories. Borrowing from the structure of scientific classification, we propose a unifying family of theories with three context-specific species (King, 2011). Species 1 involves use in a single setting (e.g., one school district or one nonprofit). Much of the existing research, which focuses on use within individual organizations, belongs to this species (e.g., Alkin et al., 1979; King & Pechman, 1984). Species 2 involves use across multiple sites, for example, within a network of organizations or across all sites that share a common funding source (e.g., Toal, Johnson, King, & Lawrenz, 2008). Species 3 is much broader in scope, involving use across time and space; it is evaluation influence, accommodating the intangible and indirect effects of the evaluation process and its results (e.g., Rebolloso, Baltasar, & Canton, 2005; Oliver, 2008). This use family is grounded in an ecological model of contexts ranging from small to large. One key caveat is the need to consider program size in each species. Evaluation use research must acknowledge the effects that the size of a program has, regardless of context, as larger programs are unavoidably more complicated and/or complex, affecting the likelihood and potential of both use and misuse.

Building in part on the framing proposed in Henry and Mark (2003) and Mark and Henry (2004), we further suggest that the field considers a realist evaluation approach for future studies, that is, identifying what works for whom in what specific contexts (Pawson, 2008; Pawson & Tilley, 1997, 2004). Realist evaluation attends to the complexity of the multiple systems within which programs work by documenting the multiple possibilities of specific contexts (C), focusing on the mechanisms (M) at work in the setting, and measuring the outcomes (O), in this case instances of use (including what has been called influence), hence context-specific CMOs. If evaluation use scholars consistently develop CMOs in their studies across species, especially using common outcome measures, then the important features of use contexts and of the mechanisms that result in use in them may finally make the content of a unifying theory evident.

At the end of Alkin and Vo (2018), his revised introductory textbook, Alkin borrows a term from Miller (2010) and presents the "theoretical signature" of a descriptive theory he calls

context-sensitive evaluation that is grounded in research on evaluation use. It indicates context (C), applicable situations, in two ways—type of evaluation (formative) and size of program (local, small scale); it identifies mechanisms (M), operational activities that the evaluator should do in order for the action to be classified as a context-sensitive evaluation; and it specifies an outcome (O), namely, use. Studying context-sensitive evaluation by compiling multiple CMOs may make it possible to develop formal research-grounded descriptive theories of evaluation use. Such research could certainly build on what we have learned from existing research. Variables to be considered include those that Kirkhart's theory highlights (time, source, and intention) plus Alkin and Taut's term *awareness*, the three levels (individual, interpersonal, and collective, i.e., public and private organizations) that Henry and Mark name, multiple subject areas (public health, social justice, education, etc.), the social context and technological advances connecting people, and so on. We believe that describing and documenting context in careful detail is imperative.

Finally, following our review of the history of evaluation use research and thoughts on how to structure its future, we want to suggest two broad topics that in our opinion represent fertile ground for additional research:

- The engagement and evaluative education of potential users: Research reviews have pointed to the importance of engaging people in evaluation processes, but multiple questions remain. What exactly does involvement or engagement mean in relation to use, and how might you measure it? Who exactly needs to be involved? What is the role of those with decision-making authority or those who champion evaluation? In addition, the growing literature on ECB suggests that evaluators may require different competencies to actively involve stakeholders in evaluations over time and teach them the skills of evaluative thinking. A central question would be: To what extent is use built in in organizations with a high level of evaluation capacity? What is the role of both internal and external evaluators in creating systems that foster routine use?
- Evaluator competence and commitment to fostering use: Factors related to an evaluator's explicit skills at developing and sustaining their clients' commitment to use represent a second area ripe for study. Such research could focus on a number of specific evaluator attributes and competencies we know to be important, for example, a full personal commitment to use, high-level interpersonal skills, facilitation and instructional skills, and an ability to develop trust and meaningful relationships in culturally appropriate ways.

This final section has included a critique of past research, a quick overview of current research, and thoughts on future research directions. Brandon and Singh (2009) wrote that the results of existing research might help thoughtful evaluation researchers to frame better studies, knowing what variables to consider and focusing more on validity issues. We agree. Even though after 50 years research on evaluation use may appear to remain in its initial stages, the pattern theory suggests undeniably that it has provided valuable insights. We would argue that the lack of common understandings to build upon and the complexity of the context have hampered this research, but that now, with this empirical grounding, greater progress can occur.

Conclusion

The earliest discussions of evaluation use held a symbolic mirror up to evaluation practice, and some saw a disappointing image. If decision makers were not using evaluation results, why spend the money to conduct program evaluations? A broadened definition and multiple research studies clarified the use picture as people came to see it as a multifaceted process dependent on many factors, some within an evaluator's control and others related to the settings where evaluations took

place. The three articles in this series (Alkin & King, 2016, 2017; current article) have recorded what we now know about evaluation use. What is the current status of evaluation use? Researchers continue to study this critically important topic, and some are applying concepts from other disciplines, expanding the theoretical frames available for understanding this complex process. We are increasingly aware of the importance of context and the likely need to develop context-specific descriptive theories. Knowing this, we conclude our historical review with hope for the future.

Acknowledgment

Declaration of Conflicting Interests

Funding

Notes

1. The third edition of the standards adds a fifth domain: evaluation accountability.
2. Recent discussion on AEA's EvalTalk blog has raised concerns that "the evaluation tree—while sacred to the dominant evaluation community—is inaccurate at best and emblematic of colonialism and capitalism at worst in that it does not recognize the role that trees play in nurturing life, as part of larger ecosystems" (Shanker, 2018, para. 1). It is not our intention here to critique the tree, but rather to apply it as it is the most common category scheme for evaluation theories.
3. In a revision of Alkin's theory tree, Mertens and Wilson (2012) add a fourth branch to identify evaluation theories that address issues related to social justice.
4. While this article is not about evaluation theory in general, we note the different ways that scholars have thought about the challenge. What might an appropriate theory be for a transdiscipline like evaluation? One possibility from Smith (2010) would create a unique and restricted definition of theory for evaluation, one that explains "underlying fundamental issues," in contrast to "theories of science, which provide empirically testable predictions." Henry and Mark (2003) and Mark and Henry (2004), among others, propose using existing theories from other disciplines. The question of whether there is or can or should be a unique and distinctive evaluation theory remains.
5. Patton is a prolific writer whose thoughts on evaluation practice continue to evolve. For the purpose of this analysis, we rely on the explicit steps that describe utilization-focused evaluation practice in Patton (2012). These steps predate Patton's more recent thinking on developmental evaluation (Patton, 2011; Patton, McKegg, & Wehipeihana, 2016), principles-focused evaluation (Patton, 2018), and blue marble evaluation.
6. Peck and Gorzalski (2009) present a third evaluation use framework, but we have excluded it here for two reasons: (a) its theoretical grounding rests on a literature review that reportedly identified "only two empirical studies" (p. 143), failing to identify the 41 additional studies discussed in the comprehensive literature review of Johnson et al. (2009) and (b) it incorrectly conflates the terms *use* and *influence*.
7. The complete list of theorists included Jennifer Greene, J. Bradley Cousins and Leithwood, Marvin Alkin, Michael Quinn Patton, J. E. Wollenberg (who had conducted dissertation research in California school districts), W. P. Johnston, Michael Huberman, K. W. Johnson, and John Owen.
8. Ottoson and Martinez (2010) describe the three cross-context variables as follows:

(1) Types of evaluation use: In analyzing their interview data, they identified seven types of evaluation use. Three of these reflect the traditional categories in the literature: instrumental, conceptual, and symbolic. Two other categories—valuing use and decision-making use—in our opinion fit within the category of instrumental use, and one category—evaluation learning use—appears to rename process use. The seventh type—communication use—for us would be an example of the traditional category of persuasive use.

(2) Threads of use: Ottoson and Martinez created this term to name "sequences among types of evaluation use" (p. 8), that is, "patterns or links among categories" (p. 15). Threads of use trace specific actions in the use process. As they summarize, "[T]hrough evaluation learning and dissemination, decisions are made, program changes occur, and funding may be attracted" (p. 15).

(3) Leveraged use: In Ottoson and Martinez's definition, this is the " . . . evaluation-related learning, tools and datasets, networks, relationships and other by-products of multiple evaluation stakeholders that influence not only the *immediate program*, but are carried across time, settings, fields, jobs and disciplines with influence elsewhere" (p. 8, emphasis in original). Similar to the concept of evaluation influence (discussed in the following section), the idea is that people take components of the evaluation experience and its outcomes and use them to advantage in settings beyond the program that was evaluated.

9. *Social betterment* is an unfortunate term for the ultimate outcome of evaluation because one person's betterment may well be another's impoverishment (King, 2015). The Nazi's final solution, for example, surely sought to create their version of a better society, but millions died as they implemented that vision. The same is true of the government agents who removed Native American children from their homes and communities to attend Bureau of Indian Affairs boarding schools.

10. It is interesting to note that the 2017 winner of AEA's Research on Evaluation Award was Michael Quinn Patton, a gifted theorist and practitioner who does not routinely conduct formal research studies. The award's explanation includes the following descriptors: "Work will have increased the empirical knowledge of the factors and conditions that influence (a) evaluation practice or (b) the understanding of evaluation theory and practice. This empirical knowledge may include, but not be limited to, exemplars for how research can make evaluation more useful, exemplars for how research can improve evaluation practice, the links between evaluation theory and practice, development of new approaches to evaluation, and the understanding of evaluation as a method to gather credible evidence" (AEA, 2018c).

11. Speaking to the global span of this research subject, evaluation scholars from throughout the world are publishing in journals in English. In all likelihood, additional studies are published in non-English evaluation or other social science journals.

References

Akanbang, B. A. A., Darko, R., & Atengdem, P. (2013). Programme implementers' experiences of process use types in three evaluation contexts in northern Ghana. *Operant Subjectivity*, *36*, 270–292.

Alkin, M. C. (1990). *Debates on evaluation*. Thousand Oaks, CA: Sage.

Alkin, M. C. (Ed.). (2013). *Evaluation roots: A wider perspective of theorists' views and influences*. Thousand Oaks, CA: Sage.

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

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

Alkin, M. C., & King, J. A. (2017). Definitions and factors associated with evaluation use and misuse. *American Journal of Evaluation*, *38*, 434–450.

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

Alkin, M. C., & Vo, A. T. (2018). *Evaluation essentials* (2nd ed.). New York, NY: Guilford Press.

American Evaluation Association. (2018a). *American evaluation association evaluator competencies*. Washington, DC: Author. Retrieved from www.eval.org

American Evaluation Association. (2018b). *American evaluation association draft revision of guiding principles for evaluators*. Washington, DC: Author. Retrieved from www.eval.org

American Evaluation Association. (2018c). *American evaluation association research on evaluation award*. Washington, DC: Author. Retrieved from www.eval.org

Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. *The Academy of Management Review*, *14*, 496–515.

Balthasar, A. (2009). Institutional design and utilization of evaluation: A contribution to a theory of evaluation influence based on Swiss experience. *Evaluation Review*, *33*, 226–256.

Baptiste, L. J. C. (2010). *Process use across evaluation approaches: An application of Q methodology in program evaluation* (Unpublished doctoral dissertation). Kent State University, Kent, OH.

Blewden, M. (2010). Developing evaluation capacity and use in the New Zealand philanthropic sector: What can be learnt from the US experience? *Evaluation Journal of Australasia*, *10*, 8.

Brandon, P. R., & Singh, J. M. (2009). The strengths of the methodological warrants for the findings on research on program evaluation use. *American Journal of Evaluation*, *30*, 123–157.

Burr, E. M. (2009). *Evaluation use and influence among project directors of state GEAR UP grants* (Doctoral dissertation). University of Tennessee, Knoxville.

Chelimsky, E. (2013). Evaluation purposes, perspectives, and practice. In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. 267–282). Los Angeles, CA: Sage.

Chen, H. T. (2015). *Practical program evaluation: Theory-driven evaluation and the integrated evaluation perspective* (2nd ed.). Los Angeles, CA: Sage.

Contandriopoulos, D., & Brousselle, A. (2012). Evaluation models and evaluation use. *Evaluation*, *18*, 61–77.

Cornachione, E. B., Trombetta, M. R., & Nova, S. P. C. (2010). Evaluation use and involvement of internal stakeholders: The case of a new non-degree online program in Brazil. *Studies in Educational Evaluation*, *36*, 69–81.

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.

Cousins, J. B., Goh, S. C., Elliott, C. J., Aubry, T., & Gilbert, N. (2014). Government and voluntary sector differences in organizational capacity to do and use evaluation. *Evaluation and Program Planning*, *44*, 1–13.

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

De Laat, B., & Williams, K. (2014). Evaluation use within the European commission: Lessons for the evaluation commissioner. In M. L. Loud & J. Mayne (Eds.), *Enhancing evaluation use: Insights from internal evaluation units* (pp. 147–174). Thousand Oaks, CA: Sage.

Donaldson, S. I., & Lipsey, M. W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. Shaw, J. Greene, & M. Mark (Eds.), *The handbook of evaluation: Policies, programs, and practices* (pp. 56–75). London, England: Sage.

Fleischer, D. N., & Christie, C. A. (2009). Evaluation use: Results from a survey of U.S. American evaluation association members. *American Journal of Evaluation*, *30*, 158–175.

Fleming, M. A. (2011). Attitudes, persuasion, and social influence: Applying social psychology to increase evaluation use. In M. M. Mark, S. I. Donaldson, & B. Campbell (Eds.), *Social psychology and evaluation* (pp. 212–243). New York, NY: Guilford.

Greenseid, L. O., & Lawrenz, F. (2011). Using citation analysis methods to assess the influence of science, technology, engineering, and mathematics education evaluations. *American Journal of Evaluation*, *32*, 392–407.

Gudmundsson, H., & Sørensen, C. H. (2013). Some use—Little influence? On the roles of indicators in European sustainable transport policy. *Ecological Indicators*, *35*, 43–51.

Hartz, Z. M., Denis, J., Moreira, E., & Matida, A. (2009). From knowledge to action: Challenges and opportunities for increasing the use of evaluation in health promotion policies and practices. In L. Potvin & D. V. McQueen (Eds.), *Health promotion evaluation practices in the Americas: Values and research practices from the Americas* (pp. 101–120). New York, NY: Springer.

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.

Højlund, S. (2014). Evaluation use in the organizational context: Changing focus to improve theory. *Evaluation*, *20*, 26–43.

Jacobson, C., Carter, R. W., Hockings, M., & Kelman, J. (2011). Maximizing conservation evaluation utilization. *Evaluation*, *17*, 1–19.

Johnson, J. B. (1998). Toward a theoretical model of evaluation utilization. *Evaluation and Program Planning*, *21*, 93–110.

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.

Karan, L. (2009). *Evaluation use in non-governmental organizations* (Dissertation). Fletcher School of Law and Diplomacy, Tufts University, Medford, MA.

King, J. A. (2011). *Practical implications of the genus and species of evaluation use*. Paper presented at a forum honoring Marvin Alkin. University of California at Los Angeles, Los Angeles, CA.

King, J. A. (2015). *"Taking* what *action for social change? Dare evaluation create a new social order?"* Keynote address, Visitors Studies Association, Indianapolis, IN.

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

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

Kirkhart, K. E. (2000). Reconceptualizing evaluation use: An integrated theory of influence: The expanding scope of evaluation use. *New Directions for Evaluation*, *88*, 5–23.

Lamm, A. J., Israel, G. D., & Harder, A. (2011). Getting to the bottom line: How using evaluation results to enhance extension programs can lead to greater levels of accountability. *Journal of Agricultural Education*, *52*, 44–55.

Lehtonen, M. (2013). The non-use and influence of UK energy sector indicators. *Ecological Indicators*, *35*, 24–34.

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

Liket, K. C., Rey-Garcia, M., & Maas, K. E. (2014). Why aren't evaluations working and what to do about it: A framework for negotiating meaningful evaluation in nonprofits. *American Journal of Evaluation*, *35*, 171–188.

Lincoln, Y. S., & Guba, E. G. (2004). The roots of fourth generation evaluation. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 225–241). Thousand Oaks, CA: Sage.

Loud, M. L., & Mayne, J. (Eds.). (2013). *Enhancing evaluation use: Insights from internal evaluation units*. Los Angeles, CA: Sage.

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

Mertens, D. M. (2013). Social transformation and evaluation. In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. 229–240). Thousand Oaks, CA: Sage.

Mertens, D. M., & Wilson, A. T. (2012). *Program evaluation theory and practice: A comprehensive guide*. New York, NY: Guilford Press.

Miller, R. L. (2010). Developing standards for empirical examinations of evaluation theory. *American Journal of Evaluation*, *31*, 390–399.

Neuman, A., Shahor, N., Shina, I., Sarid, A., & Saar, Z. (2013). Evaluation utilization research—Developing a theory and putting it to use. *Evaluation and Program Planning*, *36*, 64–70.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge, United Kingdom: Cambridge University.

Nunneley, R. D. (2010). *Causes, reasons, explanations: Three dangers of evaluation theorizing while under the influence* (Unpublished manuscript). University of Minnesota, Minneapolis.

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

Oliver, M. L. (2008). *Evaluation of emergency response: Humanitarian aid agencies and evaluation influence* (Unpublished doctoral dissertation). Georgia State University, Atlanta, GA.

Ottoson, J., & Martinez, D. (2010). *An ecological understanding of evaluation use*. Princeton, NJ: Robert Wood Johnson Foundation. Retrieved from https://www.rwjf.org/content/dam/web-assets/2010/10/an-ecological-understanding-of-evaluation-use

Patton, M. Q. (1978). *Utilization-focused evaluation*. Thousand Oaks, CA: Sage.

Patton, M. Q. (1986). *Utilization-focused evaluation* (2nd ed.). Thousand Oaks, CA: Sage.

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

Patton, M. Q. (1997). *Utilization-focused evaluation* (3rd ed.). Thousand Oaks, CA: Sage.

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

Patton, M. Q. (2011). *Developmental evaluation*. New York, NY: Guilford Press.

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

Patton, M. Q. (2018). *Principles-focused evaluation*: The guide. New York, NY: Guilford Press.

Patton, M. Q., McKegg, K., & Wehipeihana, N. (2016). *Developmental evaluation exemplars: Principles in practice*. New York, NY: Guilford Press.

Pawson, R. (2008). *The science of evaluation: A realist manifesto*. Los Angeles, CA: Sage.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Thousand Oaks, CA: Sage.

Pawson, R., & Tilley, N. (2004). *Realist evaluation, magenta book*. Strategy Unit Cabinet Office. Retrieved from http://www.communitymatters.com.au/RE_chapter.pdf

Peck, L., & Gorzalski, L. (2009). An evaluation use framework and empirical assessment. *Journal of Multi-Disciplinary Evaluation*, *6*, 139–156.

Picciotto, R. (2016). Evaluation and bureaucracy: The tricky rectangle. *Evaluation*, *22*, 1–11.

Rebolloso, E., Baltasar, F., & Canton, P. (2005). The influence of evaluation on changing management systems in educational institutions. *Evaluation*, *11*, 463–479.

Rebora, G., & Turri, M. (2011). Critical factors in the use of evaluation in Italian universities. *Higher Education*, *61*, 531–544.

Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.

Shanker, V. (2018, June 19). *Objectivity/Truth/Facts [Blog post]*. Retrieved from https://www.eval.org/p/cm/ld/fid=75

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

Smith, N. L. (2010). Characterizing the evaluand in evaluating theory. *American Journal of Evaluation*, *31*, 383–389.

Stufflebeam, D. L. (2013). The CIPP model: Status, origin, development, use, and theory. In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. 243–260). Los Angeles, CA: Sage.

Stufflebeam, D. L., & Coryn, C. L. S. (2014). *Evaluation theory, models, and applications* (2nd ed.). San Francisco, CA: Jossey-Bass.

Sturges, K. M. (2015). Complicity revisited: Balancing stakeholder input and roles in evaluation use. *American Journal of Evaluation*, *36*, 461–469.

Tennant, S. (2010). Maximising evaluation influence in an international development donor agency. *Evaluation Journal of Australasia*, *10*, 11.

Thompson, B., & King, J. A. (1981). *Evaluation utilization: A literature review and research agenda*. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles (ERIC Document Reproduction Service No. ED 199 271).

Thompson, E. H., & Patrizi, P. (2010). *Necessary and not sufficient: The state of evaluation use in foundations*. Philadelphia, PA: Evaluation Roundtable.

Toal, S. A., King, J. A., Johnson, K., & Lawrenz, F. (2008). The unique character of involvement in multi-site evaluation settings. *Evaluation and Program Planning*, *32*, 91–98.

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

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

Wholey, J. S. (2013). Using evaluation to improve program performance and results. In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (pp. 261–266). Los Angeles, CA: Sage.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2010). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

Yeary, K. H., Klos, L. A., & Linnan, L. (2012). The examination of process evaluation use in church-based health interventions: A systematic review. *Health Promotion Practice*, *13*, 524–534.