# Depicting the logic of three evaluation theories ☆

Mark Hansen [a,*], Marvin C. Alkin [a], Tanner LeBaron Wallace [b]

[a] Social Research Methodology Division, Department of Education, Graduate School of Education & Information Studies, University of California, Los Angeles, United States
[b] Psychology in Education, School of Education, University of Pittsburgh, United States

## ARTICLE INFO

## ABSTRACT

Here, we describe the development of logic models depicting three theories of evaluation practice: Practical Participatory (Cousins & Whitmore, 1998), Values-engaged (Greene, 2005a, 2005b), and Emergent Realist (Mark et al., 1998). We begin with a discussion of evaluation theory and the particular theories that were chosen for our analysis. We then outline the steps involved in constructing the models. The theoretical prescriptions and claims represented here follow a logic model template developed at the University Wisconsin-Extension (Taylor-Powell & Henert, 2008), which also closely aligns with Mark's (2008) framework for research on evaluation.

The work presented in this paper is motivated by a desire to understand the similarities and differences between various evaluation approaches or theories. As previously argued by Shadish (1998), we believe that such an understanding has value for practitioners, theorists, and those who study evaluation. Our basic premise is that visual depictions of such approaches may help to clarify their most important features, in the same way that logic models are frequently used to explicate program theories. Once the salient features are identified, it becomes more evident what particular combination of evaluator and stakeholder activities constitutes the approach. In addition, an underlying logic is revealed, including an explanation of the ways in which those activities might relate to particular effects or consequences. This may help practitioners who, for example, desire to follow a particular theorist's prescriptions. Miller (2010) has suggested that

one way in which we might assess the quality or usefulness of evaluation theories is on the basis of their operational specificity. It is well known that constructing a program logic model can sometimes reveal a less-than-fully-formed program logic. There may be similar value in depicting evaluation theories, as such efforts could help identify those aspects of a theory that are in need of further development. Finally, we believe that creating these visual representations may aid research on evaluation. The models we present here include suggestions and claims about the settings or circumstances in which an approach might be followed, the activities that should be implemented, and the effects that may be expected. From these suggestions and claims, it is possible to identify countless questions that could be examined in future empirical research. Indeed, comparing the existing research base with the assertions made in these models could inform the development of a research agenda, in which certain questions are prioritized on the basis of their relevance to theories of practice. This would allow these theories of evaluation to be tested and, over time, refined.

## 1. Background

### 1.1. A (borrowed) framework for modeling evaluation theories

As an organizing framework for our visual depictions of the three evaluation theories, we adapted a logic model template developed by the Program Development and Evaluation Unit at the University of Wisconsin—Extension (UWEX). Details of this model are described in a logic model training manual by Taylor-Powell and Henert (2008). Its core elements—including activities and outcomes—resemble those of other common logic model frameworks

(e.g., United Way, 1996; W.K. Kellogg Foundation, 2001). However, this model also attends to underlying assumptions, the situation or context of implementation, and external factors that may influence effectiveness. Our models, then, depict five basic elements. It is noteworthy that three of these elements (context, activities, and consequences, in addition to professional issues) have also been identified by Mark (2008) as potential objects of research on evaluation. Based on both the UWEX framework (Taylor-Powell & Henert, 2008) and on Mark's (2008) discussion of these categories, then, we used the following definitions:

- **Assumptions:** the underlying beliefs that inform or motivate the proposed evaluation approach.
- **Evaluation context:** the circumstances within which evaluation occurs; may include situation, needs, priorities, resources, and requirements; may include characteristics of the evaluator, the organization, and the community.
- **Evaluation activities:** the procedures used in planning and carrying out an evaluation and in communicating or disseminating findings; may include actions of both the evaluator and other stakeholders.
- **Evaluation consequences/effects:** changes that do (or do not) occur as a result of evaluation and how it was conducted; may include intended and unintended effects/changes for individuals, groups, organizations, communities, or systems.
- **External factors:** conditions or events outside the influence of the evaluation but which may affect the extent to which the evaluation accomplishes its intended effects.

In Fig. 1, we present a simplified and slightly adapted version of the UWEX template in order to illustrate how the five elements in our evaluation theory models might be expected to relate to one another. Arrows are used to suggest influence or dependence.

We should point out that representing aspects of evaluation theories in this basic form of a logic model is not an entirely new endeavor. Cousins (2003) depicted the connections between the evaluation context and policy setting, participatory evaluation practices, and various evaluation consequences (see his Figure 1). Mark and Henry (2004) elaborated on Cousins' work to visually represent a general theory of evaluation influence (see their Figure 2). Both of these earlier efforts described mechanisms or pathways in somewhat general terms—not directly tied to a particular evaluation approach. In that sense, they resemble the model in Fig. 1 (though without explicit reference to underlying assumptions). In addition, both Cousins (2003) and Mark and Henry (2004) articulate the particular characteristics of context and process that may be influential, as well as various types of effects that an evaluation could achieve. They identify, for example, the time and resources available for conducting an evaluation as important aspects of the evaluation context. We have used these previous models as guides for our work, paying close attention to those aspects of context, activities, and consequences that these authors identified. Specifically, we have incorporated those aspects as subcategories in our coding manual, described below. The primary departure in our work from these previous model-building efforts is that we attempt to describe the specific characteristics of the model elements 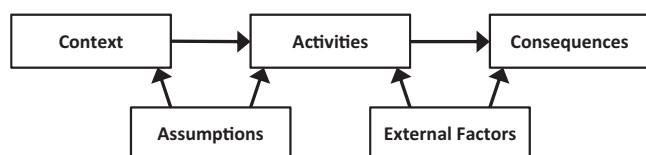that characterize three particular theories, rather than general mechanisms that may operate across many evaluation approaches.

*1.2. Evaluation approach, ideology, and theory*

With various perspectives of what constitutes an evaluation theory—and, in particular, what distinguishes such a theory from a mere approach—it is perhaps important that we clarify how we intend to use these terms. In our view, the general model presented in Fig. 1 is quite consistent with the particular view of evaluation theory described by Shadish, Cook, and Leviton (1991):

> The ideal (never achievable) theory would describe and justify why certain evaluation practices lead to particular kinds of results across situations that evaluators confront. It would (a) clarify the activities, processes, and goals of evaluation; (b) explicate relationships among evaluative activities and the processes and goals they facilitate; and (c) empirically test propositions to identify and address those that conflict with research or other critically appraised knowledge about evaluation (pp. 30–31).

Although we often use *theory* and *approach* (or *model*, for that matter) interchangeably, it would perhaps be more accurate to refer to approaches as prescriptive evaluation theories. That is, in terms of the set of activities or practices that a theorist or group of theorists prescribe and believe as the best ways to go about conducting evaluation. By this logic, we could say that the evaluation approach is nested within theory and that all good evaluation theories should include a clearly articulated approach. However, based on Shadish et al.'s (1991) criteria, evaluation theory has some additional requirements—namely testable claims or propositions about (1) when and where that approach can or should be applied (i.e., context, external factors) and (2) the results of following the approach (i.e., consequences).

Moreover, there may be certain beliefs or values (i.e., assumptions) informing and justifying the approach. Some of these may be testable, but often they are not (take, for example, beliefs about the ultimate purpose of evaluation). Smith (2007) has described such beliefs, when disconnected from context and held by a group of individuals, as evaluation *ideology*, distinct from approach or theory. However, it is not necessary that assumptions be disconnected in this way. Instead, an evaluation ideology could also be nested within evaluation theory, comprised of those fundamental assumptions and the aspects of the approach that are justified solely by those assumptions (and not by contextual considerations or desired outcomes).

In short, we believe that it is correct to speak of various alternative approaches to evaluation, and each theory we examine certainly contains an approach. However, we also find that there are often specific and testable claims or propositions attached to an approach. In such cases, it is appropriate to consider the approach, assumptions, and claims as together constituting a theory of evaluation. Moreover, the structure of a logic model provides a reasonable framework for depicting how these elements come together as theory.

## 2. Methods

*2.1. Selection of evaluation theories*

Our basic approach was to identify texts from the evaluation literature that described specific theories of evaluation. We examined the content of those texts and created visual summaries that took the familiar form of a logic model. One of our hopes in developing visual depictions of evaluation theories was the



**Fig. 1.** A General model for the logic of an evaluation theory.

possibility that such representations could facilitate comparisons between approaches. Accordingly, we elected to examine, in depth, three theories that we expected would include both common and distinctive elements for this study. With their "Evaluation Theory Tree," Alkin and Christie (2004, Christie & Alkin, 2008, 2012) categorized theories on the basis of their primary emphasis: *use*, *methods*, or *valuing*. By selecting one theory (or theorist) from each branch, we hoped to apply our method to a relatively diverse group. At the same time, we intentionally chose theories that have been articulated fairly recently and, as a result, were likely to be influenced by and incorporate ideas from other theories (e.g., Alkin, 1991). Thus, we aimed for approaches that would vary in emphasis but also contain some overlap in content. Each of the theories we considered is associated with a specific author or group of co-authors and is described in sufficient detail by the author(s) to be viewed as a coherent approach. Ultimately, we settled upon three theories. From the *use* "branch" of Christie and Alkin's (2008) "Theory Tree," we selected Cousins' practical participatory evaluation (PPE; e.g., Cousins and Whitmore, 1998); from the *values* "branch," Greene's values-engaged approach (VEE; e.g., Greene, 2005a, 2005b); and from the *methods* "branch," Mark, Henry, and Julnes' emergent realist evaluation (ERE; e.g., Mark, Henry, & Julnes, 1998).

### 2.2. Selection of data sources

For each theory, we searched for publications by the corresponding authors using web-based search engines, as well as reference lists from texts in which the authors were cited. Our goal was to identify a small number of publications that would collectively constitute a clear and reasonably complete exposition of the theory. Journal articles, book chapters, books, and conference papers were all considered. Writings by the authors that did not clearly describe the particular approach or theory of interest were excluded from our analysis. At the same time, it was not required that the particular name or label we now attach to the theory be mentioned. The terms "value-engaged" and "values-engaged," for example, appear in only three of the publications we considered (Greene, 2005a, 2005b; Greene, DeStefano, Burgon, & Hall, 2006). However, other writings by Greene present ideas that are clearly consistent with—and, in our view, may reasonably be considered as articulations of—VEE. Thus, we have included these in our analysis, as they contribute to our understanding of this theory. After identifying candidate materials and conducting some initial coding (described in the following section), we contacted each of the theorists, providing a brief description of this study and our proposed reference list. We asked for suggestions regarding any additional published materials that might help us understand their evaluation theory. We then screened each suggested work and added those found to include relevant information to our lists. The final reference lists are presented in Tables 1–3.

We limited our analyses to the writings of those we most closely associated with each theory. This means that there may be

**Table 1**
Data sources: Practical participatory evaluation (PPE).

Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14*, 397–418.
Cousins, J. B. (1996). Consequences of researcher involvement in participatory evaluation. *Studies in Educational Evaluation, 22*, 3–27.
Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation, 80*, 5–23.
Cousins, J. B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 245–266). Boston, MA: Kluwer.
Cousins, J. B., & Earl, L. M. (1999). When the boat gets missed: Response to M. F. Smith. *American Journal of Evaluation, 20*, 309–317.

**Table 2**
Data sources: Values-engaged evaluation (VEE).

Greene, J. C. (1997). Evaluation as advocacy. *Evaluation Practice, 18*, 25–35.
Greene, J. C. (2001). Evaluation extrapolations. *American Journal of Evaluation, 22*, 397–402.
Greene, J. C. (2005a). A value-engaged approach for evaluating the Bunche-Da Vinci Learning Academy. *New Directions for Evaluation, 106*, 27–45.
Greene, J. C. (2005b). Evaluators as stewards of the public good. In S. Hood, R. K. Hopson, & H. T. Frierson (Eds.), *The role of culture and cultural context: A mandate for inclusion, truth, and understanding in evaluation theory and practice* (pp. 7–20). Greenwich, CT: Information Age Publishing.
Greene, J. C., DeStefano, L., Burgon, H., & Hall, J. (2006). An educative, values-engaged approach to evaluating STEM educational programs. *New Directions for Evaluation, 109*, 53–71.
Greene, J. C., Millett, R. A., & Hopson, R. H. (2004). Evaluation as a democratizing practice. In M. T. Braverman, N. A. Constantine, & J. K. Slater (Eds.), *Foundations and evaluation: Contexts and practices for effective philanthropy* (pp. 96–118). San Francisco, CA: Jossey-Bass.
Ryan, K., Greene, J. C., Lincoln, Y., Mathison, S., & Mertens, D. M. (1998). Advantages and challenges of using inclusive evaluation approaches in evaluation practice. *American Journal of Evaluation, 19*, 101–122.

**Table 3**
Data sources: Emergent realist evaluation (ERE).

Henry, G. T. (2000, October). Using evaluation findings for policy: A realist perspective. *Paper presented at the European Evaluation Society Conference*. Lausanne, Switzerland.
Henry, G. T., & Julnes, G. (1998). Values in realist evaluation. *New Directions for Evaluation, 78*, 53–71.
Mark, M. M., Henry, G. T., & Julnes, G. (1999). Toward an Integrative Framework for Evaluation Practice. *American Journal of Evaluation, 20*, 177–198.
Mark, M. M., Henry, G. T., & Julnes, G. (1998). A realist theory of evaluation practice. *New Directions for Evaluation, 78*, 3–32.
Mark, M. M., Henry, G. T., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. San Francisco, CA: Jossey-Bass.

other authors who have written about a particular theory or closely related approaches. There is a rich literature concerning various forms of participatory evaluation, for example. Though we exclude such writings in this analysis, we recognize that such writers may indeed have made substantial contributions to these theories and the ways in which they are practiced. However, focusing on a primary author (or group of authors, in the case of ERE) allows us to better understand one particularly important and influential conception of the theory. In addition, this approach allows us to verify our representations with these primary theorists, who provide their feedback and critiques later in this volume. It would certainly be worthwhile to also examine alternative conceptions—and the ways in which those add to or alter the primary authors' basic theories. Such comparisons, however, are beyond the scope of this study.

### 2.3. Coding of data sources

Within the five elements, we identified sixteen sub-categories or dimensions and created codes for 61 distinct concepts, drawing from Cousins (2003), Mark and Henry (2004), and Mark (2008). These concepts—and their organization by element and sub-category—are listed in the Appendix. The coding of evaluation texts was completed by a group of five graduate students, including the first author of this paper and the authors of the three papers that follow. After developing and reviewing the coding manual, the group pilot tested the manual by independently coding a single text. The results were compared, and discrepancies in coding were discussed and resolved. The codebook was then refined in order to clarify the appropriate assignment of each code. Analysis of the 17 texts selected for the study then proceeded with each text being

carefully read and analyzed by one reviewer. Excerpts that were assigned one or more codes were copied into a database. A second reviewer then examined the excerpts and their corresponding codes to check their accuracy. This reviewer then updated the database with any necessary changes.

Upon completion of the coding process, the excerpts from the texts for each evaluation theory were sorted by category and used to guide the development of the logic models. Specifically, we tried to identify the major ideas conveyed by the authors with respect to each of the logic model elements: assumptions, context, activities, consequences, and external factors. Our goal was to produce models that would be acceptable simplifications of these approaches, conveying the most important features clearly and accurately. We began by sorting the excerpts for each theory by category. This allowed us to see all the excerpts for which we assigned codes related to evaluation context, for example. By reading these excerpts, we identified the major themes and wrote summary statements. These statements were added to the appropriate fields of the logic models. This approach was intended to limit the content of the models to statements that could be directly supported in the texts we examined, reducing the possibility that the models might reflect our own ideas and not those of the evaluation theorists. That said, there are places where we used some discretion—specifically, in the placement and ordering of the statements in the logic models. In many cases, the texts we consulted did not provide a strong basis for determining the nature of the relationships (such as temporal or causal) between statements in our depictions. So in the same way that an evaluator might suggest ways of organizing the activities and goals identified by stakeholders when developing a program logic model, so we too have, out of necessity, imposed an ordering and logic to the concepts described in these theories.

## 3. Results

### 3.1. Observed code frequencies

Table 4 presents a summary of the codes assigned to excerpts from the texts chosen for each evaluation theory. The number of codes assigned ranged from 238 for PPE to 419 for ERE; 336 codes were used for VEE. Although frequencies may serve as a rather crude measure of relative emphasis, we do observe some trends that seem consistent with our understanding of the core ideas of each theory. For example, the source texts for PPE included a relatively large number of references to evaluation consequences, especially use of findings and increases in organizational capacity. In addition, we assigned codes related to underlying assumptions (particularly, concerning epistemology and the ultimate purpose of evaluation) much more frequently with ERE than for the other two theories. More in-depth discussion of three of the most emphasized elements of the logic models—context (Vo, 2013), activities (Dillman, 2013), and consequences (Luskin & Ho, 2013)—appear in the papers that follow.

### 3.2. Logic models for evaluation theories

The logic models developed for the three evaluation approaches are shown in Figs. 2–4. These depictions represent our attempt to summarize the key ideas of each theory. As in many common logic model frameworks, these models may be read from left to right. The context statements describe the setting and circumstances in which the evaluation activities are implemented, including conditions that exist prior to the evaluation. The activities, in turn, may be viewed as a response to the context or situation. We indicate this potential influence of context on evaluation activities with arrows. Activities may also reflect (or be the result of) certain

**Table 4**
Code frequencies, by category and evaluation theory.

| Category | PPE | | VEE | | ERE | |
|---|---|---|---|---|---|---|
| Subcategory | Count | % | Count | % | Count | % |
| **Assumptions** | | | | | | |
| Program | 0 | – | 2 | 0.6 | 2 | 0.5 |
| Epistemology | 1 | 0.4 | 6 | 1.8 | 29 | 6.9 |
| Evaluation | 5 | 2.1 | 42 | 12.5 | 61 | 14.6 |
| Stakeholders | 1 | 0.4 | 3 | 0.9 | 14 | 3.3 |
| Other | 0 | – | 4 | 1.2 | 6 | 1.4 |
| Assumptions subtotal | 7 | 2.9 | 57 | 17.0 | 112 | 26.7 |
| **Context** | | | | | | |
| Evaluator | 31 | 13.0 | 39 | 11.6 | 36 | 8.6 |
| Organization/program | 23 | 9.7 | 7 | 2.1 | 8 | 1.9 |
| Stakeholders | 16 | 6.7 | 6 | 1.8 | 56 | 13.4 |
| Other | 9 | 3.8 | 3 | 0.9 | 4 | 1.0 |
| Context subtotal | 79 | 33.2 | 55 | 16.4 | 104 | 24.8 |
| **Activities** | | | | | | |
| Planning/preparation | 24 | 10.1 | 44 | 13.1 | 44 | 10.5 |
| Conducting the evaluation | 47 | 19.7 | 97 | 28.9 | 77 | 18.4 |
| Reporting | 0 | – | 8 | 2.4 | 12 | 2.9 |
| Activities subtotal | 71 | 29.8 | 149 | 44.3 | 133 | 31.7 |
| **Consequences/effects** | | | | | | |
| Characteristics of evaluation | 9 | 3.8 | 20 | 6.0 | 7 | 1.7 |
| Perceptions of evaluator | 1 | 0.4 | 4 | 1.2 | 1 | 0.2 |
| Organizational capacity | 37 | 15.5 | 5 | 1.5 | 15 | 3.6 |
| Process use | 6 | 2.5 | 9 | 2.7 | 2 | 0.5 |
| Use of findings | 23 | 9.7 | 21 | 6.3 | 31 | 7.4 |
| Other consequences | 2 | 0.8 | 11 | 3.3 | 12 | 2.9 |
| Consequences subtotal | 78 | 32.8 | 70 | 20.8 | 68 | 16.2 |
| **External factors** | | | | | | |
| External factors subtotal | 3 | 1.3 | 5 | 1.5 | 2 | 0.5 |
| **Total (all categories)** | 238 | | 336 | | 419 | |

Note: For the *Assumptions* category, subcategories identified here refer to the specific concepts/codes.

assumptions on the part of the theorists concerning the ultimate purpose(s) of evaluation or the particular ways in which an evaluation ought to be done.

Consequences are connected to the evaluation activities and to each other by arrows. These arrows are used to indicate possible (hypothesized) causal relationships. For example, in the model for PPE (Fig. 2), the ultimate intended consequence is increased program effectiveness. It is proposed that this outcome is a result of improvements in organizational decision-making and problem-solving (i.e., *if* decision-making and problem-solving are improved, *then* program effectiveness is increased), which may depend on other effects (specifically, the use of research and evaluation findings in decision-making, the validity of findings, and shared understanding of program functions and processes) first being achieved.

Evaluation consequences are believed to be affected by both the evaluation activities and external factors. As it turned out, we came across very little discussion of such factors. However, according to all three theories, values and support of stakeholders and decision-makers may be expected to affect the potency of the evaluation. For VEE and ERE, values influence the extent to which evaluation findings receive attention and are used in decision-making. For PPE, support and engagement may moderate the effect of the evaluation process on the development of organizational capacity.

## 4. Discussion

We have described the development of logic models representing three theories of evaluation. There are, of course, several limitations in the approach we have taken. First, we have sampled a small number of texts from a limited set of authors. Although these texts were chosen with the hope that they would serve as representative expositions of the theories, it is possible that selection of additional sources would result in changes to the

**Context**      **Activities**      **Consequences/Effects**

| | |
|---|---|
| **Evaluator…**<br>• has knowledge of evaluation logic and methods (technical skills)<br>• is capable of training others in evaluation<br>• is tolerant of imperfection and mistakes | Evaluator trains stakeholders in skills necessary to implement the evaluation |

Stakeholders develop technical knowledge, research skills (i.e., capacity for systematic inquiry)

Stakeholders develop shared understanding of program functions and processes

**Participant stakeholders…**
• are decision-makers (managers, developers, sponsors, implementers)
• are willing to learn
• are committed to organizational learning
• need information to support decision making and problem solving

Selected participants are involved in all aspects of the evaluation: planning, data collection, data analysis, interpretation, formulating judgments

Development of dense interpersonal networks

Evaluation is integrated into organizational culture

Evaluation is perceived as credible

Stakeholders use evaluation and research findings in decision making

Evaluation meets the information needs of stakeholders (responsive, relevant, timely)

Improved decision making and problem solving

**Program…**
• is still in development
• has clear goals that stakeholders generally agree upon

Evaluator assumes responsibility for technical evaluation tasks (maintains technical rigor)

Evaluation findings are valid

Increased program effectiveness

**Assumptions**          **External Factors**

• Knowledge is socially constructed
• Stakeholders have the ability to learn research skills, given appropriate training
• Integrating evaluation into the organization takes time and several repetitions

• Micropolitics and the level of ongoing organizational support may influence the success of the participatory approach

**Fig. 2.** Logic model for practical participatory evaluation (PPE).

models. In particular, it is possible that elements of the logic model that were rarely discussed in the texts we examined would be more fully described elsewhere. For example, we identified very little treatment of external factors.

It is also likely that, in attempting to represent only those concepts that were explicitly conveyed in the texts, we failed to capture implicit concepts. It is possible, for example, that theorists may take some ideas as a given and, therefore, unnecessary to articulate. The theorists included in this study are certainly aware of accepted professional standards of practice, such as the American Evaluation Association's *Guiding Principles for Evaluators* (Newman, Scheirer, Shadish, & Wye, 1995) and the *Program*
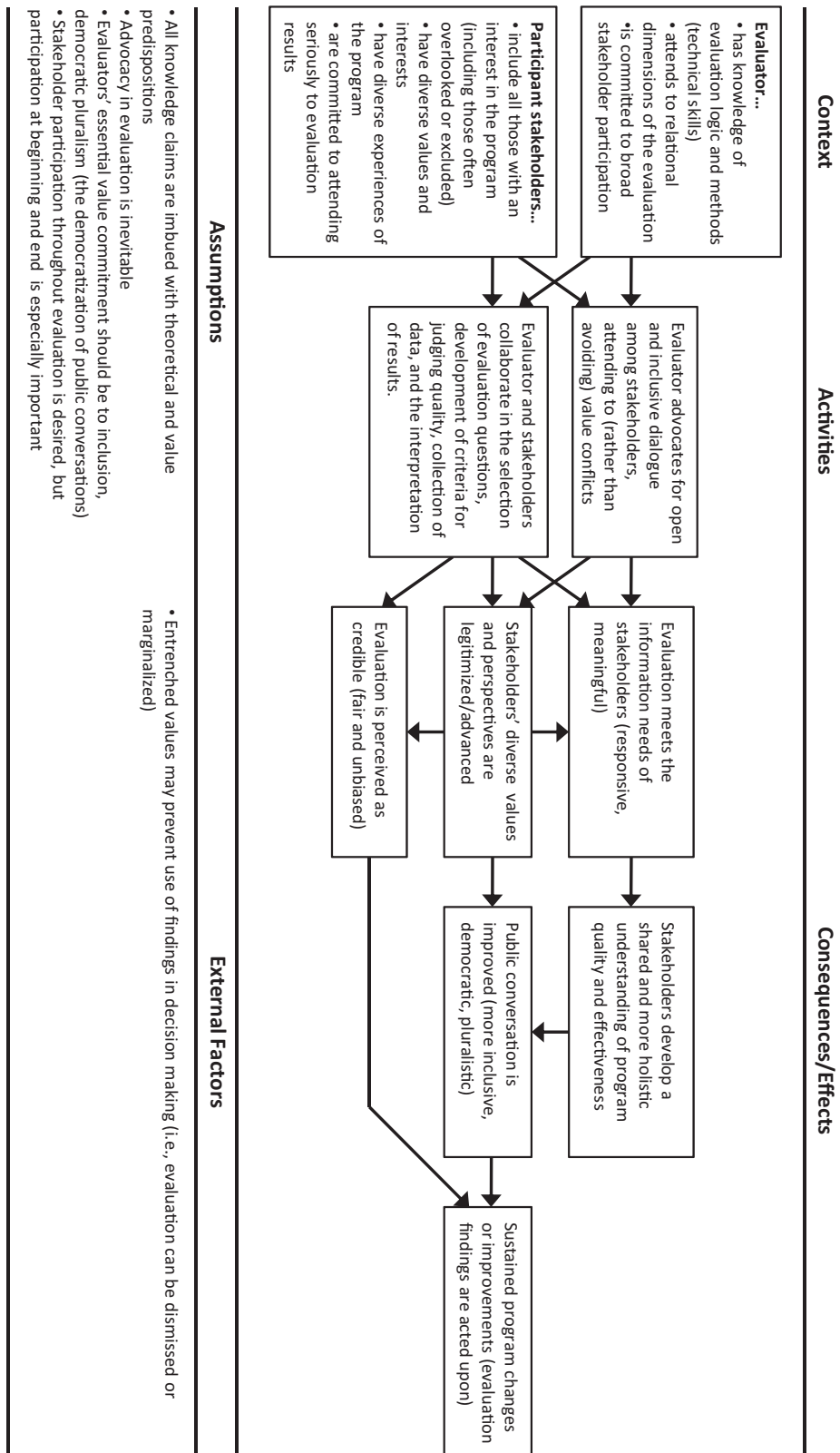
**Fig. 3.** Logic model for values-engaged evaluation (VEE).

*Evaluation Standards* (Joint Committee on Standards for Educational Evaluation, 1999). It is possible, then, that there are certain aspects of each of these theories that are assumed by the authors but left unstated. The approach that we used did not readily lend itself to representing these ideas.

It should also be noted that theorists' views may change over time. We have treated all the texts as being equally useful in understanding the authors' theoretical perspectives. However, it is possible that more recent selections may better reflect their current views. In the extreme case, it is possible that perspectives

**Context**  **Activities**  **Consequences/Effects**

**Program or Policy Context**
• Alternative programs or policies are being considered (i.e., a "fork in the road")
• There are opportunities for shifts in funding options
• Program or policy enjoys solid support and is unlikely to be substantially altered
- OR -
• The program or policy is newly established

**Evaluator…**
• recognizes contextuality of program effects
• acknowledges strengths and weaknesses of various modes of inquiry and researchmethods

**Stakeholders…**
• include all potential users, including the public and democratically established institutions
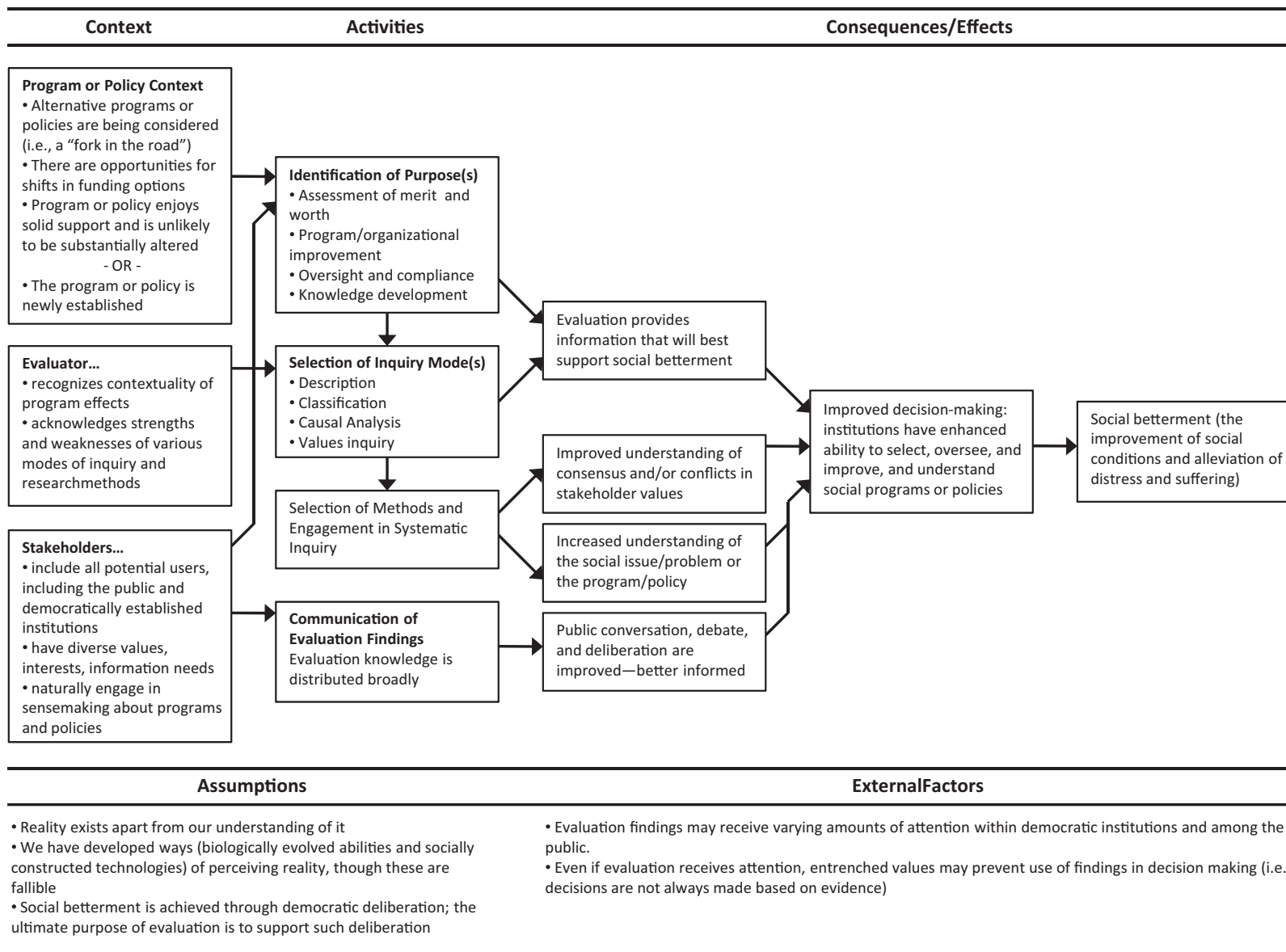• have diverse values, interests, information needs
• naturally engage in sensemaking about programs and policies

**Identification of Purpose(s)**
• Assessment of merit and worth
• Program/organizational improvement
• Oversight and compliance
• Knowledge development

**Selection of Inquiry Mode(s)**
• Description
• Classification
• Causal Analysis
• Values inquiry

Selection of Methods and Engagement in Systematic Inquiry

**Communication of Evaluation Findings**
Evaluation knowledge is distributed broadly

Evaluation provides information that will best support social betterment

Improved understanding of consensus and/or conflicts in stakeholder values

Increased understanding of the social issue/problem or the program/policy

Public conversation, debate, and deliberation are improved—better informed

Improved decision-making: institutions have enhanced ability to select, oversee, and improve, and understand social programs or policies

Social betterment (the improvement of social conditions and alleviation of distress and suffering)

**Assumptions**  **ExternalFactors**

• Reality exists apart from our understanding of it
• We have developed ways (biologically evolved abilities and socially constructed technologies) of perceiving reality, though these are fallible
• Social betterment is achieved through democratic deliberation; the ultimate purpose of evaluation is to support such deliberation

• Evaluation findings may receive varying amounts of attention within democratic institutions and among the public.
• Even if evaluation receives attention, entrenched values may prevent use of findings in decision making (i.e., decisions are not always made based on evidence)

**Fig. 4.** Logic model for emergent realist evaluation (ERE).

at different points in time should be viewed as entirely different theories. Building models based on writings across many years may thus lead to the combining of ideas that were actually intended to describe different approaches.

An inherent problem in logic modeling is the fact that there are always many ways of representing a program (or, in this case, an evaluation theory). No representation is perfect, of course. However, some may be decidedly inconsistent with the views or intentions of the developer (theorist). Leeuw (2003) noted that such "mis-reconstructions" of theory, if used as guides for practice or decision-making, could be dangerous. It is important to note, therefore, that we do not advocate the use of these models for such purposes, nor would we suggest that one could gain a complete understanding of these evaluation theories without studying the primary sources, including those we reviewed in order to develop the models. Our modest goal, then, was only to summarize complex ideas in ways that could be easily conveyed and compared.

In developing a program logic model, it is typical (and, arguably, crucial) to involve stakeholders, including program developers (e.g., McLaughlin & Jordan, 1999). Even if the evaluator assumes primary responsibility for drawing the model, he or she may present early drafts to stakeholders for their feedback. This provides a way of ensuring the accuracy of the models—essentially asking the question, "Have I understood you correctly?" In this study, we have consulted the theorists' writings, but not the theorists themselves. As a result, we do not consider the models we present here as being final. Later papers in this volume include the theorists' responses to our efforts. In that way, this volume serves as a written record of the sorts of conversations that often occur in the course of depicting a program's logic. We share the models as they are in the hope that they may provoke further discussion and with the expectation that they will change as a result of those conversations.

Finally, we recognize that the decision to represent these evaluation theories as logic models introduces a particular structure—one that is basically linear, among other things—that some readers (and theorists) may question. This choice may limit the extent to which our depictions realistically capture how these theories are experienced in practice. That said, our goal was to simplify and reduce complex ideas to their key elements. That inevitably means some loss of information. As statistician George E.P. Box explained, "Essentially, all models are wrong, but some are useful" (Box & Draper, 1987, p. 424). Our hope, then, is that our depictions are not so wrong as to have become useless.

## 5. Conclusion

We have described the development of logic models of three evaluation theories. These models were constructed based on the writings of the authors most closely associated with each theory. After careful reading and analysis of selected texts, we wrote statements to summarize the major themes related to the basic assumptions, contextual considerations, prescribed activities, intended consequences, and influential external factors discussed in each theory. These statements were arranged within the familiar form of a program logic model. Although we note many limitations in this approach, we believe that these depictions may contribute to our understanding of these theories and facilitate comparisons between them. The remaining papers in this collection will examine the concepts and claims we have depicted, thereby illustrating some of the possible uses for these models.

## Appendix A

Summary of codes used in analysis of evaluation theory texts
See Tables A1–A5.

**Table A1**
Assumptions: the underlying beliefs that inform or motivate the proposed evaluation approach (5 concepts/codes).

| Sub-category | Concept | Definition/guiding questions |
|---|---|---|
| General | Program | What are the assumptions concerning programs that motivate this approach? |
| | Epistemology | What are the assumptions concerning knowledge or reality underlying the approach? |
| | Evaluation | What are the assumptions concerning evaluation (its purpose, goals, etc.) that motivate this approach? |
| | Stakeholders | What are the assumptions concerning stakeholders that motivate this approach? |
| | Other | What are other assumptions that motivate this approach? |

**Table A2**
Evaluation context: the circumstances within which evaluation occurs; may include situation, needs, priorities, resources, and requirements; may include characteristics of the evaluator, the organization, and the community (20 concepts/codes).

| Sub-category | Concept | Definition/guiding questions |
|---|---|---|
| Evaluator | Technical expertise | What sorts of technical or methodological expertise/training should the evaluator have in order to implement this evaluation approach? |
| | Content knowledge | What sorts of substantive knowledge should the evaluator have in order to implement this evaluation approach? |
| | Interpersonal skills | What sorts of interpersonal skills should the evaluator possess in order to implement this evaluation approach? |
| | Competencies | What sorts of knowledge, skills, or competencies should the evaluator possess in order to implement this evaluation approach? |
| | Values | What would be the values, beliefs, priorities, orientation, and attitudes of the evaluator who pursues this evaluation approach? |
| | Other | Other characteristics of the evaluator that are needed in order to implement this evaluation approach. |
| Organization/program | Program characteristics | For what kind of program would this evaluation approach be appropriate? For what stage of program development/implementation? What would be true of the program or its design/underlying theory? What characteristics of the program would favor (or not) implementation of this evaluation approach? |
| | Administrative support | What resources/assistance (including administrative support) from the organization/program is needed in order to implement this evaluation approach? |
| | Values | What sort of organizational values, priorities, interests, or culture would favor (or not) the implementation of this evaluation approach? |

**Table A2** (Continued)

| Sub-category | Concept | Definition/guiding questions |
|---|---|---|
| | Information needs | What are the information needs of the organization/program that the evaluation approach is (or is not) well-suited to address? |
| | Skills | What (and whose) skills, knowledge, and expertise are required within the program/organization in order to implement this evaluation approach? |
| | Other | Other aspects of the *organizational context* in which the evaluation takes place. |
| Stakeholders | Identity | Who are the individuals and groups (in addition to evaluator and program/organization staff) who have a stake in the evaluation? Who might be considered an audience for the evaluation? |
| | Values | What are the values, priorities, and interests of those individuals and groups (in addition to evaluator and program/organization staff) who have a stake in the evaluation? |
| | Information needs | What are the information needs of stakeholders that the evaluation approach is (or is not) well-suited to address? |
| General | Impetus | What events, conditions, or circumstances lead to the evaluation? |
| | Time | Are there time constraints that would make this evaluation approach particularly suitable (or unsuitable)? |
| | Resources | What general resources (e.g., money) are necessary in order to implement this evaluation approach? |
| | Politics | What sorts of political circumstances (local or otherwise) would favor (or not) the implementation of this evaluation approach? |
| | Other | Other aspects of the *general context* in which the evaluation takes place. |

**Table A3**

Evaluation activities: the procedures used in planning and carrying out an evaluation and communicating or disseminating findings; may include actions of both the evaluator and other stakeholders (16 concepts/codes).

| Sub-category | Concept | Definition/guiding questions |
|---|---|---|
| Planning/preparation | Stakeholder selection | What stakeholders should be involved in the evaluation? How should these stakeholders be identified or chosen? |
| | Stakeholder participation | What is the role of stakeholders in planning or designing the evaluation? Who should be involved, and in what ways? |
| | Understanding the program | What activities are implemented to help the evaluator or stakeholders understand the program? |
| | Designing the evaluation | What steps are involved in designing the evaluation (including decisions about study questions, methods, etc.)? |
| | Other | What other activities are implemented in planning the evaluation? |
| Conducting the evaluation | Evaluator's role | What is the role of the evaluator in conducting the evaluation? What are her/his responsibilities? |
| | Stakeholder participation | In what ways (and to what extent) should stakeholders be involved in the evaluation? What are their roles and responsibilities? |
| | Study design | What study design(s) are typical of this approach? |
| | Data collection | What's the approach to data collection? Are particular sources of data or collection methods emphasized? Who is involved in collection? |
| | Data analysis | What's the approach to data analysis? Are there particular methods that are typical of the approach? Who is involved data analysis? |
| | Developing conclusions | How are conclusions formed under this approach? What is the basis for a conclusion? Whose is it? |
| | Other | What other activities are involved in conducting the evaluation? |
| Reporting | Report generation | How should information about the evaluation be communicated to stakeholders? What information? With whom should it be shared? |
| | Developing recommendations | Does the evaluation result in the formation of recommendations? If so, who is involved in developing those recommendations? And what is their nature (e.g., to whom are they addressed)? |
| | Other | What other activities are involved in communicating the findings of the evaluation? |
| Outputs | Outputs | What measures should be used to assess the way in which the approach was implemented? |

**Table A4**

Evaluation consequences/effects: changes that do (or do not) occur as a result of the evaluation and how it was conducted; may include intended and unintended effects/changes for individuals, groups, organizations, communities, or systems (19 concepts/codes).

| Sub-category | Concept | Definition/guiding questions |
|---|---|---|
| Characteristics the evaluation | Responsiveness | Evaluation is (perceived by stakeholders as being) more responsive or relevant (and timely) to the needs of stakeholders. |
| | Findings credibility | Evaluation findings are (perceived by stakeholders as being) credible. |
| | Validity | Evaluation findings are (more) accurate, valid (including internal and external, etc.). |
| | Other | Other characteristics or perceptions of the evaluation (specifically, those thought to be due to the approach that was implemented). |
| Perceptions of the evaluator | Evaluator credibility | Evaluator establishes rapport, trust with stakeholders; evaluator is viewed by stakeholders as credible. |
| | Other | Other perceptions of the evaluator (specifically, those thought to be due to the approach that was implemented). |
| Organizational capacity | Systematic inquiry | Does the evaluation approach enhance the capacity of (or enthusiasm for) the organization to engage in future evaluation work or other forms of systematic inquiry? |
| | Communication | Does the evaluation approach enhance the organization's ability to communicate the value of the program to stakeholders? |
| | Knowledge | What knowledge does the organization acquire through the evaluation? |
| | Decision-making | Does the evaluation approach enhance organizational decision-making? |
| | Program effectiveness | Does the evaluation approach enhance program effectiveness? |
| | Sustainability | Does the evaluation approach enhance program or organization's sustainability? |
| | Other | Are there any other ways the evaluation approach enhances the organizational capacity? |

**Table A4** (*Continued*)

| Sub-category | Concept | Definition/guiding questions |
|---|---|---|
| Process use | Process use | In what ways might knowledge of or participation in the evaluation affect stakeholders? (Note: may overlap with other concepts.) |
| Use of findings | Conceptual | Are the evaluation findings used for informative/educative purposes (e.g., to shape others' views about the program)? |
| | Instrumental | Are the evaluation findings used to make specific decisions? |
| | Symbolic | Are the evaluation findings used to persuade or legitimize? |
| | Other | Are there other ways in which the findings are used? |
| Other consequences | Other | Are there other changes that result from the evaluation? |

**Table A5**
External factors: Conditions (external to the evaluation itself, beyond its influence) that could affect the extent to which the evaluation achieves (or does not achieve) its intended effects (1 concept/code).

| Sub-category | Concept | Definition/guiding questions |
|---|---|---|
| General | External factors | What conditions outside the direct influence of the evaluation or control of the evaluator (including circumstances, events, attitudes, beliefs, actions, etc.) may influence evaluation effects or consequences? |

## References[1]

Alkin, M. C. (1991). Evaluation theory development: II. In M. W. McLaughlin & D. C. Phillips (Eds.), *Evaluation and education: At quarter century* (pp. 91–112). Chicago, IL: University of Chicago Press.

Alkin, M. C., & Christie, C. A. (2004). An evaluation theory tree. In M. C. Alkin (Ed.), *Evaluation roots* (pp. 12–65). Thousand Oaks, CA: SAGE Publications.

Box, G. E. P., & Draper, N. R. (1987). *Empirical model-building and response surfaces*. Oxford: Wiley.

Christie, C. A., & Alkin, M. C. (2008). Evaluation theory tree re-examined. *Studies in Educational Evaluation, 34,* 131–135.

Christie, C. A., & Alkin, M. C. (2012). An evaluation theory tree. In M. C. Alkin (Ed.), *Evaluation roots* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

*Cousins, J. B. (1996). Consequences of researcher involvement in participatory evaluation. *Studies in Educational Evaluation, 22,* 3–27.

*Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14,* 397–418.

*Cousins, J. B., & Earl, L. M. (1999). When the boat gets missed: Response to M. F. Smith. *American Journal of Evaluation, 20,* 309–317.

Cousins, J. B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan & D. Stufflebeam (Eds.), *International Handbook of Educational Evaluation* (pp. 245–266). Boston, MA: Klewer.

*Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation, 80,* 5–23.

Dillman, L. (2013). Comparing evaluation activities across multiple theories of practice. *Evaluation and Program Planning, 38,* 53–60.

*Greene, J. C. (1997). Evaluation as advocacy. *Evaluation Practice, 18,* 25–35.

*Greene, J. C. (2001). Evaluation extrapolations. *American Journal of Evaluation, 22,* 397–402.

*Greene, J. C. (2005a). A value-engaged approach for evaluating the Bunche-Da Vinci Learning Academy. *New Directions for Evaluation, 106,* 27–45.

Greene, J. C. (2005b). Evaluators as stewards of the public good. In S. Hood, R. K. Hopson, & H. T. Frierson (Eds.), *The role of culture and cultural context: A mandate for inclusion, truth, and understanding in evaluation theory and practice* (pp. 7–20). Greenwich, CT: Information Age Publishing.

Greene, J. C., DeStefano, L., Burgon, H., & Hall, J. (2006). An educative, values-engaged approach to evaluating STEM educational programs. *New Directions for Evaluation, 109,* 53–71.

Greene, J. C., Millett, R. A., & Hopson, R. H. (2004). Evaluation as a democratizing practice. In M. T. Braverman, N. A. Constantine, & J. K. Slater (Eds.), *Foundations and evaluation: Contexts and practices for effective philanthropy* (pp. 96–118). San Francisco, CA: Jossey-Bass.

*Henry, G. T. (2000). Using evaluation findings for policy: A realist perspective. *Paper presented at the European Evaluation Society Conference.*

*Henry, G. T., & Julnes, G. (1998). Values in realist evaluation. *New Directions for Evaluation, 78,* 53–71.

Joint Committee on Standard for Educational Evaluation. (1999). *The Program Evaluation Standards* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Leeuw, F. L. (2003). Reconstructing program theories: Methods available and problems to be solved. *American Journal of Evaluation, 24,* 5–20.

Luskin, R. J. C., & Ho, T. (2013). Comparing the intended consequences of three theories of evaluation. *Evaluation and Program Planning, 38,* 61–66.

Mark, M. M. (2008). Building a better evidence base for evaluation theory: Beyond general calls to a framework of types of research on evaluation. In N. L. Smith & P. Brandon (Eds.), *Fundamental issues in evaluation* (pp. 111–134). New York, NY: Guilford Press.

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

*Mark, M. M., Henry, G. T., & Julnes, G. (1998). A realist theory of evaluation practice. *New Directions for Evaluation, 78,* 3–32.

*Mark, M. M., Henry, G. T., & Julnes, G. (1999). Toward an integrative framework for evaluation practice. *American Journal of Evaluation, 20,* 177–198.

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

McLaughlin, J. A., & Jordan, G. B. (1999). Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning, 22,* 65–72.

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

*Ryan, K., Greene, J. C., Lincoln, Y., Mathison, S., & Mertens, D. M. (1998). Advantages and challenges of using inclusive evaluation approaches in evaluation practice. *American Journal of Evaluation, 19,* 101–122.

Shadish, W. R., Cook, T. D., & Leviton, L. A. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage. pp. 36–67.

Newman, D., Scheirer, M. A., Shadish, W. R., & Wye, C. (Eds.). (1995). The American Evaluation Association's guiding principles. *New Directions in Evaluation, 66,* 19–26.

Shadish, W. R. (1998). Evaluation theory is who we are. *American Journal of Evaluation, 19,* 1–19.

Smith, N. L. (2007). Empowerment evaluation as evaluation ideology. *American Journal of Evaluation, 28,* 169–178.

Taylor-Powell, E., & Henert, E. (2008). *Developing a logic model: Teaching and training guide*. Madison: University of Wisconsin-Extension. Retrieved from http://www.uwex.edu/ces/pdande/evaluation/pdf/lmguidecomplete.pdf.

United Way of America. (1996). *Measuring program outcomes: A practical approach*. Arlington, VA: United Way of America.

Vo, A. T. (2013). Visualizing context through theory deconstruction: A content analysis of three bodies of evaluation theory literature. *Evaluation and Program Planning, 38,* 44–52.

Wallace, T. L., & Alkin, M. C. (2007). Identifying critical processes and outcomes across evaluation approaches. *Paper presented at the annual conference of the American Evaluation Association.*

W.K. Kellogg Foundation. (2001). *Logic model development guide: Logic models to bring together planning, evaluation & action*. Battle Creek, MI: WK Kellogg Foundation.

**Mark Hansen**, MPH, is a graduate student in the Graduate School of Education & Information Studies at the University of California, Los Angeles. His research interests include educational measurement and the evaluation of educational and social programs.

**Marvin C. Alkin**, EdD, is professor emeritus of Social Research Methodology in the Graduate School of Education & Information Studies at the University of California, Los Angeles. His research focuses on evaluation utilization, evaluation theory, and problems of evaluating educational programs. He is the editor of Evaluation Roots: Tracing Theorists' Views and Influences and author of Evaluation Essentials: From A to Z.

**Tanner LeBaron Wallace**, PhD, is an assistant professor of Applied Developmental Psychology at the University of Pittsburgh. Her research interests include identifying the most effective mechanisms for promoting positive youth development, exploring how schools and youth programs serve as important contexts for youth development, and supporting the knowledge development function of evaluation studies.

---

[1] * indicates text used as data source for model development.