

Evaluation theory tree re-examined

Christina A. Christie^{a,*}, Marvin C. Alkin^b

^aSchool of Behavioral and Organizational Sciences, Claremont Graduate University, 123 East Eighth Street, Claremont, CA 91711, United States

^bUniversity of California, Los Angeles, United States

ARTICLE INFO

Keywords:
Evaluation
Evaluation theory

ABSTRACT

When examining various evaluation prescriptive theories comparatively, we find it helpful to have a framework showing how they are related that highlights features that distinguish theoretical perspectives, thus a “theory” about theories. The evaluation theory tree that we presented in Alkin’s recent book, *Evaluation Roots* [Eisner, E. (2004). *The roots of connoisseurship and criticism: A personal journey*. In M. Alkin (Ed.), *Evaluation roots: Tracing theorists’ views and influences*. Thousand Oaks, CA: Sage; Guba, E., & Lincoln, Y., (2004). *The roots of fourth generation evaluation: Theoretical and methodological origins*. In M. Alkin (Ed.), *Evaluation roots: Tracing theorists’ views and influences*. Thousand Oaks, CA: Sage; Stake, R. (2004). *Stake and responsive evaluation*. In M. Alkin (Ed.), *Evaluation roots: Tracing theorists’ views and influences*. Thousand Oaks, CA: Sage; Wholey, J. S. (2004). *Using evaluation to improve performance and support policy decision making*. In M. Alkin (Ed.), *Evaluation roots: Tracing theorists’ views and influences*. Thousand Oaks, CA: Sage], is such a framework. Just as theorists modify their views over time, in this paper we suggest modifications to the theory tree presented in the *Roots* book, including a repositioning of a few theorists, the addition of theorists, and a reconceptualization of the valuing branch.

© 2008 Elsevier Ltd. All rights reserved.

Introduction

As we began thinking about this paper we were reminded of a quotation by Egon Guba that Alkin has used many times. The setting was a meeting of the American Educational Research Association where Alkin’s students role played different theorists, indicating how they (as that theorist) would modify their point of view based upon ideas two Belgian philosophers put forth. Guba, as a discussant to the student paper reflecting his views, said:

You see, in many ways I am not the “real” Egon Guba—at least not the one that Lindheim had in mind when she conducted her exercise. That Egon Guba was the one working and writing seven to 10 years ago... my mind has changed about so many things that it is hard for me to recall what perspective I held then (1979, p. 139).

Guba’s statement reminds us that views change over time. People read things. Ideas get floated. Interpretations get modified. That is one way that evaluation theories develop and change over time. But, even with the possibility that changes may take place, it

is important to have an explicit conceptual framework. Such frameworks help to guide practitioners when choosing an approach to use, as well serve as valuable teaching tools by offering comparative, relational depictions of approaches.

As we think about the various evaluation prescriptive theories, we find it helpful to have a theory about how they are related—a framework of prescriptive theories. The evaluation theory tree offered in Alkin’s recent book, *Evaluation Roots* (2004), is such a system. Our view is that there are three basic elements in considering evaluation theories: use, methods, and valuing. All theorists are concerned with the methods that will be employed in conducting the evaluation. All theorists recognize that evaluation is an enterprise that involves valuing (distinguishing it from most research). All theorists recognize that evaluations will be used in ways that affect programs. We postulate, however, that theorists differ in the particular emphasis they place on one or another of these dimensions, which we refer to as “branches” of the evaluation theory tree. Thus, we can place theorists on the branch that best reflects their primary emphasis in a manner that reflects some combination of history and/or the influence of a particular approach on another (either building upon, or in response to the tenets of a particular approach).

As theorists modify their views over time, we propose changes to the evaluation theory tree (version 3), which reflect some substantive changes in our thinking. In this paper we suggest

* Corresponding author. Tel.: +1 909 607 9020; fax: +1 909 607 9009.
E-mail address: tina.christie@cgu.edu (C.A. Christie).

additional modifications of the theory tree presented in the Roots book (see Fig. 1). Fig. 2 shows the tree as we now perceive it.

Evaluation theory tree 2004: brief summary

After several years of in-depth conversation about evaluation theory and theorists, we decided to illustrate the relationship and loosely map the evolution of evaluation theory development, focusing primarily on work that emerged in North America. Our “evaluation theory tree” was the product of this effort. We must acknowledge here, as we do in the original version of the tree, that we are not talking about theories as defined by traditional academic standards. While the term *theory* is conventionally used in the evaluation literature, it is more appropriate to use the terms approaches, models, or even frameworks. Evaluation “theories” are almost exclusively prescriptive, that is, they offer a set of rules, prescriptions and prohibitions that specify what a good or effective evaluation study is and how an evaluation study should be conducted. None of the evaluation approaches is predictive or offers an empirical theory. Nevertheless, it is the convention of the evaluation literature to refer to our prescriptive approaches as theories, and so we too refer to the models described in this paper as theories and to those who have developed these models as theorists.

The first “theory tree” presented and described in great detail in Chapter 2 of Alkin’s Roots book (2004). In this chapter, Alkin and Christie discuss 27 different evaluation approaches and classify each by their primary focus on one of three essential elements of evaluation, *methods*, *values*, and *use*. The authors posit that all prescriptive theories of evaluation much consider issues related to: (1) the *methods* used in an evaluation, including the study design, (2) the manner in which data are to be judged and valued and by whom, and the underlying *values* used to accomplish this and (3) the *use* of the evaluation effort. The final chapter of the book offers an analysis of the classification schema in light of information provided by the theorists about their work in the text, and a second version of the theory tree was presented (see Fig. 1), differing only slightly from the one presented earlier.

In the last chapter of the Roots book, we already recognized changes from the tree’s original incarnation (presented in Chapter 2 of Roots), stimulated by the entries written by various theorists. These changes were relatively minor. By and large, they were simply repositioning sub-branches in different ways to reflect the extent to which a theorist on a particular branch tended towards the views reflected in another branch. Additionally, we had indicated two foundations for the theory tree: the tradition of social inquiry and the desire for accountability and control. Responding to Yvonna Lincoln’s comments we changed one of the foundations to social accountability and fiscal control.

Our evaluation theory tree consists of a trunk and three primary branches. The trunk initially rested upon the foundational roots of accountability and systematic social inquiry. The middle branch of the tree grows from the social inquiry foundational root. The primary focus of those placed on this branch is on developing models for evaluation practice which at the core are grounded in and derived from social science research methods. Theorists’ models are mostly derivations of the randomized control trail, and are intended to offer results that are generalizable and have a focus on “knowledge construction.” The valuing branch focuses on those theorists who as a central feature of their evaluation theories consider the process of placing value on the evaluation as the essential component of an evaluator’s work. This extends to include theorists who advocate for the evaluator to systematically facilitate the determining value by others. The third major branch is use. Use theorists are most concerned with the use of evaluation itself and the information generated from the evaluation is used and focuses on those who will use the information.

Evaluation theory tree 2006

Fig. 2 presents the third version of our evaluation theory tree. The modifications of both the *use* and *methods* branches are relatively minor, particularly when compared to the revision of the *value* branch. The authors invite comment and critique of the modifications presented, as we view this work as ongoing, requiring further thinking and analysis.

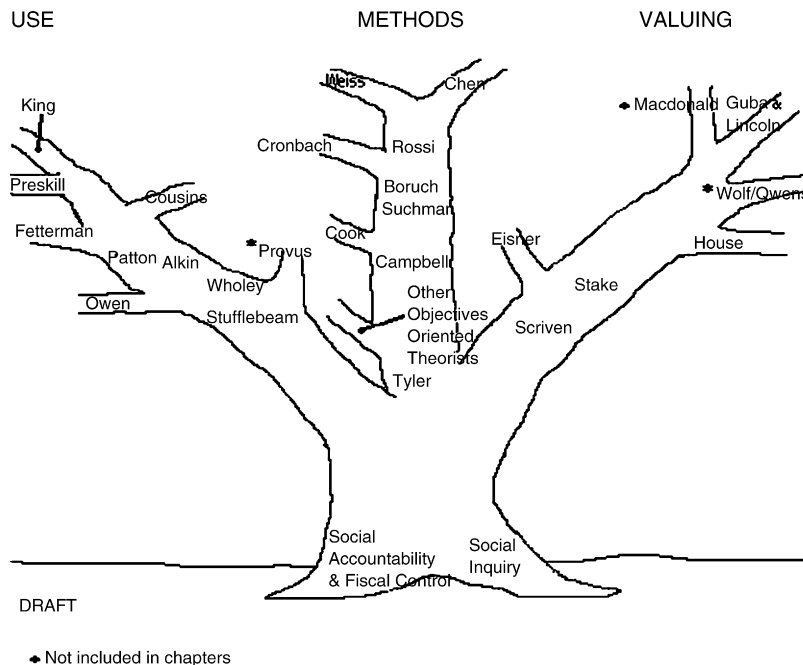


Fig. 1.

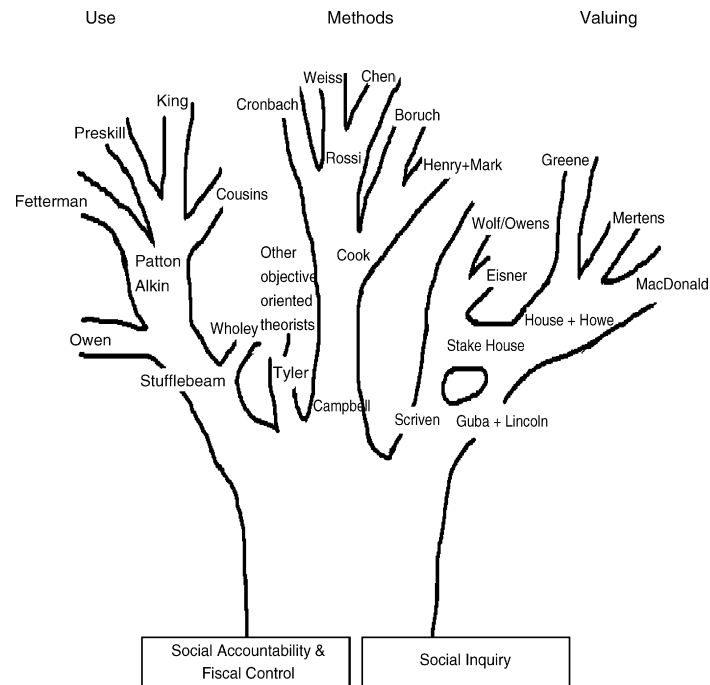


Fig. 2.

Use branch revisions

In its most recent incarnation, the Use branch underwent only minor revision and still closely resembles the Tree presented in Chapter 26 of the *Evaluation Roots* book. The slight modifications include repositioning Joseph Wholey's work (2004) to reflect an emphasis towards the Methods branch. Also, we deleted Malcolm Provus from the branch because much of his work has had relatively little influence on other perspectives on the branch.

Methods branch revisions

The Methods branch received more substantial changes. Ralph Tyler was repositioned to a sub-branch to reflect that while his theoretic point of view was in fact heavily methods related (objectives based evaluation), he was not a theoretical predecessor of those further up on the branch. His original positioning was intended to reflect his influence on the field of educational evaluation (which was very significant), but upon further reflection we concluded that his overall influence on the methods branch specifically was less than his original position implied. Thus, we placed him on a small sub-branch near the base of the branch.

Donald Campbell continues to be the heart and soul of this branch both in the way that he directly influenced other theorists with his work on experiments, quasi-experiments, and validity, as well as the way in which his views provided a basis for counterpoint, particularly with respect to Lee Cronbach's views.

We have also deleted Edward Suchman from the branch. We recognize the important influence of Suchman's writing on evaluation—particularly in positioning Campbell's work prominently in the evaluation discourse. Nonetheless, despite his historical significance, it did not seem appropriate to continue to include him on the branch because he himself did not offer the field a specific evaluation model. Rather, he promoted Campbell's work as the most effective approach for conducting evaluation studies to measure program impact. Peter Rossi continues to be

shown on the main portion of the branch as an influence on both Carol Weiss and, more particularly, on Huey Chen.

A significant change on this branch was moving Tom Cook who previously had been positioned in a sub-branch flowing from Campbell. He is now positioned on the main branch, and his sub-branch was deleted. A new sub-branch was created for Bob Boruch. Boruch's sub-branch comes directly out of Campbell and Cook. This placement reflects the steadfast influence of Campbell's randomized control trial (RCT) work on Boruch. The branch is offset on a sub-branch to emphasize his departure from both Cook and Campbell in his lack of regard for quasi-experiments. This position also offsets Boruch from the others on the methods branch, all of whom have embraced quasi-experiments as an acceptable method for studying causality, although everyone on the branch would agree that the experiment is ideal if context and conditions permit.

Another major change on this branch is the addition of Gary Henry, Mel Mark and George Julnes. In the *Roots* book we said:

(T)heories included were able to be classified onto a single branch of the tree... Whether some theories were not included due to their comprehensiveness or our conceptual inability is unclear. A particular example comes to mind: the work of Mark, Henry, and Julnes (2000). These authors view social betterment as the ultimate objective of evaluation and present a point of view grounded in what they refer to as a 'common sense realist philosophy'... The very diversified nature of this perspective, while a great strength in presenting an understanding of evaluation, precludes its inclusion on the tree (pp. 58–59).

We admit that in this instance our conceptual analysis was incomplete. Our views in determining exclusion were based heavily on the Mark et al. (2000) book. However, in reflecting further on the writings of these authors, we were further struck by the views presented in the "realist evaluation" monograph in *New Directions for Evaluation* (Henry, Julnes, & Mark, 1998). Based upon a more in-depth reading of this text, we have placed Henry, Mark, and Julnes on the methods section branch, leaning toward the post-positivist valuing branch. The leaning highlights Henry and Mark's collaboration with George Julnes to create "a new theory

that captures the sense-making contributions from post-positivism and the sensitivity to values from constructivist traditions” (Henry et al., 1998, p. 1).

Henry, Mark, and Julnes describe their approach, Emergent Realist Evaluation, as a comprehensive new evaluation model that offers re-conceptualized notions of use, methods, and valuing. Valuing and use are important aspects of Mark and Henry’s theory, however methodology is the core pillar. Emergent Realist Evaluation is (1) an evaluation methodology that gives priority to the study of generative mechanisms, (2) attentive to multiple levels of analysis, and (3) mixed methods appropriate. Mechanisms are defined as “the underlying causes of the changes that are observed” (Mathison, 2005, p. 360). Henry, Mark, and Julnes also see evaluation as a tool for social and political change within democracies. In their words, “social betterment, rather than the more popular and pervasive goal of utilization, should motivate evaluation” (Mark, Henry, & Julnes, p. 19). Based on this description, we place Henry, Mark, and Julnes on the Methods branch of the tree in a location following Tom Cook, leaning toward the valuing branch.

Valuing branch revisions

The valuing branch has received the most substantial revision. This branch has always been the most difficult to explain in relationship to its evolution. It is obvious that the work of Scriven has influenced it significantly. After all, it is Scriven who proclaims that an evaluation is not evaluation without valuing; in his words, evaluation is the science of valuing (Scriven, 2003). It is the work of the evaluator to make a value judgment about the object that is being evaluated. This thinking has shaped and defined the field. However, this branch also includes the work of those interested in social justice in evaluation, as well as those who espouse the philosophy of subjectivity, that is, the claim that there is no one objective reality. Here, theorists are also concerned with valuing, but the emphasis is on the extent to which values shape the evaluation. There are questions as to whose values should shape the evaluation, why, and with what intent. This is quite different from the type of valuing that concerns Scriven. With careful examination it has become clear how these two foci differ.

With little surprise, examining the theoretical perspectives on the valuing branch within the context of the philosophy of science offers a revised understanding of how we understand these perspectives. The basic axioms of the post-positivist and constructivist paradigms offer a clearer framework to further categorize evaluation models on the valuing branch. Thus, we have elected to split the value branch in two, naming the left arm of the branch stretching toward the methods branch as “valuing: post-positivist influence” and the right arm, “valuing: constructivist influence.” It is important to stress the use of the word *influence* in our description of the newly reshaped valuing branch. That is, some perspectives on the valuing branch are shaped more exactly by a paradigm, while others feel only a paradigm’s undercurrent.

Views of science shifted, however, during the 20th century away from positivism into post-positivism. Post-positivists recognize that all observation is fallible and has error. Where positivists believed that the goal of science was to uncover the truth, the post-positivist believes that the goal of science is to attempt to measure truth, even though that goal cannot be obtained. Constructivism is one element of interpretivism and ontologically takes a relativist stance. There is no single, tangible reality that can be approximated; there are only multiple, constructed realities. Epistemologically, constructivism views

subjectivity as the only reality, that is, the only way the unknown can become known is through our own, individual belief systems.

Scriven’s realist thinking about valuing is reflective of both the ontology and epistemology of the post-positivist paradigm. He does not argue against the idea that we should be seeking an objective truth about the object being evaluated. He, in fact, offers what he believes to be a comparatively unbiased method for obtaining truth about an object’s worth, and then advocates for the evaluator to make a value judgment after gathering the most credible evidence. Further, he does not reject the idea of using experiments to determine causality, but rather argues that there should be more than one method for determining causality (Donaldson, Christie & Mark, 2008). Scriven’s thinking pushed the field to consider valuing as a central feature of evaluation more so than anyone else. However, he has been repositioned on the valuing branch at the base of the post-positivist influence arm to reflect the ontology and epistemology of his perspective.

Stake’s work (2004) respects Scriven’s thinking, although he argues for using “thick description” to assess a program’s worth via the case study method. The use of case study methods introduced the idea that value is bound to context and that evaluators must consider context when determining value. Stake does, however, leave the valuing to the evaluator, and so does not seem to reject the realist idea that evaluation is a science of valuing. It was Stake’s inclusion of and emphasis on case study methods and program context that prompted a shift from realist to relativist inquiry models in evaluation. Thus, Stake sits on a bridge squarely between the split of the two emphases of the valuing branch.

House’s work (e.g., House, 19) brought significant attention to the inclusion of the values of the underrepresented to the evaluation process. Ontologically and epistemologically House’s work is grounded in constructivist thinking, however methodologically his work breaks from the basic axioms of the constructivist paradigm. That is, he is willing and likely to use quantitative methods when conducting evaluations. Thus, he too sits on the bridge to the right of Stake. Eisner’s work (2004) is now depicted as an offshoot from Scriven’s branch off toward the constructivist influence arm. Eisner is similar to Scriven in that he posits that the evaluator is the expert and thus determines the final value of a program. He differs from Scriven in that he believes the evaluator has the authority to judge a program merit because of one’s expert understanding of the subject area (i.e., education, public health), rather than because of one’s expertise as an evaluator. Because Eisner uses intense observation (which includes both quantitative and qualitative measures, but emphasizes qualitative), he is placed on an offshoot branch leaning toward the constructivist influence arm of the valuing branch. Wolf/Owens (Owens, 1973; Wolf, 1979) are placed on the post-positivist influence arm because their perspective does adhere to the principle that evaluators should pursue the truth.

We have moved Guba and Lincoln (2004) to the base of the constructivist influence arm to reflect the ontology and epistemology of their perspective, which closely adheres to the basic tenets of the constructive paradigm. Their theoretical model serves as the foundation for a string of models that are influenced by the tenets of the constructivist paradigm, three of which have been added to the theory tree. First, House and Howe’s deliberative democratic model (1999) has been included as its own model separate from the social justice work of House, which is already represented on the tree. Deliberative democratic evaluation is based on the ideals of House’s earlier work on social justice in evaluation and has been

influenced by the processes put forth in Guba and Lincoln's *Fourth Generation Evaluation*. Thus, deliberative democratic evaluation has been placed on the tree above Guba/Lincoln.

House and Howe believe that “evaluators should accept authority but not power” (House & Howe, 1999, p. 102). An evaluator who caters to those with power perpetuates inequality, which is why inclusion is the first criterion of deliberative democratic evaluation. Inclusion wards off stakeholder bias and involves stakeholders with and without power to participate in the evaluation. Dialogue is the next criterion. Dialogue is necessary to ensure that stakeholder contributions are well thought out and may be honestly weighed along with all other stakeholder contributions to the evaluation. Deliberation, the final criterion for deliberative democratic evaluation, is the weighing of each contribution to generate an accurate conclusion. House and Howe's approach is value-engaged. Fact and value are not mutually exclusive. They exist on a continuum where a middle ground exists between “brute fact” and “bare values” (House & Howe, 1999, p. 6). House and Howe describe the deliberative democratic evaluation process as follows: “. . . we can imagine moving along the value-fact continuum from statements of preferences and values collected through initial dialogue, through deliberations based on democratic principles, to evaluative statements of fact” (House & Howe, 1999, p. 100).

Jennifer Greene's approach has also been added to the tree, directly above House and Howe. Greene's value-engaged approach (Greene, 2005) is rooted in deliberative democratic evaluation principles and procedures; however, it places additional emphasis on framing the evaluation on stakeholders' values. Greene's value-engaged approach uses three criteria of deliberative democratic evaluation: inclusion, dialogue, and deliberation. However, she stresses stakeholder involvement, which closely resembles participatory evaluation approaches. She also explicitly emphasizes the use of mixed-methods designs and fieldwork within her evaluations.

Donna Mertens's inclusive approach (1999) could be considered a direct decedent of Guba/Lincoln but is unique in its emphasis on diversity and the inclusion of diverse groups. Mertens is best known for her inclusive/transformational model of evaluation in which the evaluator's primary role is to include marginalized groups, not to act as decision maker. Although the evaluator advocates for the inclusion of the marginalized groups, the evaluator does not advocate for the marginalized groups. Mertens maintains that evaluators working within an inclusive framework should ask themselves the following questions at the planning stages of the evaluation:

- Are we including people from both genders and diverse abilities, ages, classes, cultures, ethnicities, families, incomes, languages, locations, races, and sexualities?
- What barriers are we erecting to exclude a diversity of people?
- Have we chosen the appropriate data collection strategies for diverse groups, including providing for preferred modes of communication? (Mertens, 1999, p. 8)

The aim of the inclusion is to ensure that the evaluation is conducted within the context of the program's entirety in a way that discourages bias, not to advocate for one group over another. A primary goal of inclusive/transformational evaluation is to challenge the status quo in a quest to transform society. Mertens agrees with and quotes Chelmsky, who wrote that challenging the status quo is “our most important task and the best justification for our work” (quoted in Mertens, 1999, p. 2). In this paradigm evaluations are a tool to combat social inequality and social injustice.

Summary

And so, the above discussion of the modified theory tree represents our current thinking about how to classify evaluation theorists. This current picture of prescriptive evaluation theories will guide our thinking about evaluation issues until further changes become necessary.

References

- Donaldson, S. I., Christie, C. A., & Mark, M. (Eds.). (2008). *What Counts as Credible Evidence in Evaluation and Evidence-based Practice?* Thousand Oaks, CA: Sage.
- Eisner, E. (2004). The roots of connoisseurship and criticism: A personal journey. In M. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences*. Thousand Oaks, CA: Sage.
- Greene, J. C. (2005). A value-engaged approach for evaluating the Bunche-Da Vinci learning academy. *New Directions for Evaluation*, 106, 27–47.
- Guba, E., & Lincoln, Y. (2004). The roots of fourth generation evaluation: Theoretical and methodological origins. In M. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences*. Thousand Oaks, CA: Sage.
- Henry, G., Julnes, G. J., & Mark, M. (1998). *Realist evaluation*. San Francisco: Jossey-Bass.
- House, E., & Howe, K. (1999). *Values in evaluation and social research*. Thousand Oaks, CA: Sage.
- Mark, M., Henry, G., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. San Francisco, CA: Jossey-Bass.
- Mathison, S. (Ed.). (2005). *Encyclopedia of evaluation*. Thousand Oaks, CA: Sage.
- Mertens, D. (1999). Inclusive evaluation: Implications of transformational theory for evaluation. *American Journal of Evaluation*, 20(1), 1–14.
- Owens, T. (1973). Education evaluation by adversary proceeding. In E. R. House (Ed.), *School evaluation: The politics and process*. Berkeley, CA: McCutcheon.
- Scriven, M. (2003). Evaluation in the new millennium: The transdisciplinary view. In S. I. Donaldson & M. Scriven (Eds.), *Evaluating Social Programs and Problems: Vision for the New Millennium*. Mahwah, NJ: Erlbaum. pp. 19–42.
- Stake, R. (2004). Stake and responsive evaluation. In M. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences*. Thousand Oaks, CA: Sage.
- Wholey, J. S. (2004). Using evaluation to improve performance and support policy decision making. In M. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences*. Thousand Oaks, CA: Sage.
- Wolf, R. L. (1979). The use of judicial evaluation methods in the formulation of educational policy. *Educational Evaluation and Policy Analysis*, 1(93), 19–28.

Christina A. Christie is associate professor, co-director of the Masters of Arts Program in Psychology and Evaluation, and associate director of the Institute of Organizational and Program Evaluation Research in the School of Behavioral and Organizational Sciences at Claremont Graduate University.

Marvin C. Alkin is professor Emeritus, Social Research Methodology, Graduate School of Education and Information Studies, University of California, Los Angeles.