

2

The authors describe and compare the four theoretical perspectives represented by the theorists writing Chapters Three through Six, providing additional comparative analysis based on their published works.

What Theorists Say They Do: A Brief Description of Theorists' Approaches

Christina A. Christie, Tarek Azzam

Perhaps now more than ever before, funders are requesting that social and educational programs be evaluated for accountability and to demonstrate effectiveness. This is great news for the discipline and for those conducting evaluations because this mandate translates into, by and large, greater value for our work and an increased opportunity for evaluators to practice our craft. With increased occasion to evaluate, we are more likely to encounter greater diversity in the types of programs we are studying. How, then, does one determine the best evaluation design? The evaluation theory literature provides a myriad of approaches to help inform this decision and subsequently guide evaluation design and implementation. The question remains about how a theory will translate into practice.

The purpose of this volume is to examine, comparatively, the practical application of theorists' approaches to evaluation by examining four evaluations of the same case. The thought is that when asked to evaluate the same program (holding the case constant), the practical distinctions between theorists' approaches become evident. We would also speculate that subtler differences in approaches will be highlighted, and surprising similarities will emerge. For this exercise to prove profitable, theorists must be proponents of approaches that are, from the outset, different and that these differences will remain evident in their practice.

Who Was Invited and Why: Theorist Selection

We are fortunate that the four distinguished scholars invited to participate in our exercise—Jennifer Greene, Gary Henry, Stewart Donaldson, and Jean King—all agreed to take part. When selecting theorists for participation, the

editors were committed to inviting only theorists who are actively evaluating programs. Most of those who offer theoretical views on evaluation have conducted evaluations. We were concerned, however, that theorists with more limited or less current experience would be inclined to reiterate their theoretical writings rather than explain what they would actually *do* when confronted with designing an evaluation for our case school, Bunche–Da Vinci.

It can be difficult to translate into text some of the more nuanced aspects of a theoretical approach. Thus, another criterion used when selecting the evaluation scholars for this volume was contemporary and, more important, accessible theoretical writings. It was our hope that theorists whose writings are more easily understood by broad audiences would be more likely to translate their theoretical approaches into readable evaluation designs.

Using Taxonomies

Evaluation theory classification schemas systematically and thoughtfully group together and distinguish approaches based on an articulated set of criteria. The editors wanted to ensure, as much as possible, that differences in approaches would emerge when described in our exercise. With these criteria in mind, we turned to theoretical classification schema to guide our selection of theorists for this volume.

Evaluation theoretic approaches are often better understood when classified by defining features or characteristics. There are several well-known classification schemas, such as those put forth by Fitzpatrick, Sanders, and Worthen (2003), Shadish, Cook and Leviton (1991), and Alkin and House (1992). Alkin and Christie (2004) recently published their own classification schema, the evaluation theory tree. This taxonomy uses three principal dimensions of evaluation—methods, values, and use—to categorize approaches based on their primary emphasis.

Although only one of our four theorists (King) is represented in Alkin and Christie's published schema, we found that examining each perspective within the context of this particular taxonomy illustrates the diversity of theoretical perspectives represented by our theorists. It also provides an opportunity to consider what distinguishes each of these theorists. Our analysis places Donaldson on the methods branch, Greene on the values branch, and King on the use branch. Henry, through his writings (Mark, Henry, and Julnes, 2000), showed that he would not easily be placed in the methods, value, or use category.

Greene is a proponent of a value-engaged approach to evaluation and has been in the forefront of advancing both the discussion and application of this approach. This approach incorporates elements of responsive evaluation with principles from democratic evaluation. Value-engaged evaluation is responsive to the particulars of the evaluation context (Stake, 1980) while

including relevant interests, values, and views so that conclusions can be unbiased in value as well as factual aspects (House and Howe, 2000). The approach relies heavily on the values of those involved in the program, including program recipients.

The methods branch of the tree represents theoretical approaches that are grounded in, and have a primary emphasis on, social science methodology. Donald Campbell anchors this branch, from which all method branch theories have been derived. Henry, who generally evaluates large-scale education programs with an eye toward policy reform, espouses emergent realist evaluation (ERE) theory (Mark, Henry, and Julnes, 1998, 2000). This contingency-based approach aimed at social betterment generally lends itself to quantitative methods of inquiry, such as fully randomized experiments and quasi-experiments. Therefore, when investigating causation, Henry advocates using a randomized controlled trial whenever feasible. When experiments are not viable, Henry is most likely to turn to the less rigorous but often more practical quasi-experimental designs to study program effectiveness. In addition to Henry's writings on ERE, he has authored several noted contributions on sampling and other techniques necessary for and concerned with conducting good experiments. In addition, in the reports from studies Henry has conducted, his emphasis and his belief in the value for experimental methods is evident.

Donaldson is best known for his work in the area of theory-driven evaluation (TDE). Thus, Donaldson would be placed further up the methods branch, positioned next to Huey Chen. (In the context of the theory tree, moving up a branch reflects the evolution of theoretical approaches, away from, but still grounded in, the initial theoretical approaches and thinking that served as the foundation for, and are positioned at the base of, the branch.) Early TDE thinking was put forth by two of evaluation's most illustrious theorists, Carol Weiss and Peter Rossi (see Weiss, 1998; Rossi and Freeman, 1993), both of whom are on the methods branch of the theory tree. Their initial ideas were expounded on by Huey Chen, who published the first text dedicated exclusively to this approach (Chen, 1990) and is frequently recognized for ushering TDE into its prominent place in the evaluation theory literature. Donaldson refers to his variety of TDE as *program theory-driven evaluation science*.

King is concerned most with use of the evaluation process and its results. Her emphasis is not only on use, but also participation in the evaluation process as a means for increasing use and building internal program evaluation capacity. Her approach evolves most directly from participatory evaluation and extends the notions of Cousins and Whitmore (1998) by particularly emphasizing the importance of building evaluation capacity. Although Cousins is also concerned with building evaluation capacity through participation in the evaluation process, it is his secondary rather than his principal interest. King is concerned primarily with building evaluation capacity as a principal outcome for the evaluation. She argues that

increasing evaluation capacity is essential for promoting and ensuring evaluation use. Stakeholder participation is requisite for achieving this goal. King has written extensively about her experiences implementing her approach, interactive evaluation practice, not only as an external but also as an internal evaluator for a midwestern school district.

Brief Descriptions of the Theorists' Approaches

Jennifer C. Greene. Greene is an advocate of a value-engaged approach to evaluation, which incorporates responsive (Stake, 1980, 1987) and democratic (House and Howe, 2000) processes to study social and educational programs. Her approach is responsive in that the evaluation design is developed in response to the program context. It is democratic in that it places great importance on considering the perspectives and values of all legitimate stakeholders, especially stakeholders who are typically alienated from evaluation processes (House and Howe, 2000). Greene (2000) argues that evaluations can be “used to surface and legitimate differing views and values and move [stakeholder] towards shared understanding of the values of educational outcomes” (p. 16).

Greene offers three “justifications” for including stakeholder views when conducting evaluations: (1) pragmatic justification, (2) emancipatory justification, and (3) deliberative justification. The *pragmatic* justification argues for stakeholder inclusion because it increases the chance of evaluation use (Patton, 1997) and organizational learning (Preskill and Torres, 1999). The *emancipatory* justification focuses on the importance of acknowledging the skills and contributions of stakeholders and empowering them to be their own social change agents (Fettermen, 2001). The *deliberative* justification argues that evaluation should serve to ensure that program or policy conversations include all relevant interests and are “based on the democratic principles of fairness and equity and on democratic discourse that is dialogic and deliberative” (Greene, 2000, p. 14).

Greene’s methodical approach to the evaluation process reflects her interest in representing the values of differing stakeholders. She prefers to incorporate a mix of qualitative and quantitative methods (Greene and Caracelli, 1997), which can answer both broad and in-depth evaluation questions. Based on experiences during a previous program evaluation, Greene argues that the use of mixed methods should be defensible politically, philosophically, and technically (Greene and Caracelli, 1997). (In this evaluation, which is discussed in depth in Greene, 2000, stakeholders argued and clashed over the validity of the methods used and were not able to focus on discussions of values or reach any kind of meaningful consensus.) Her approach when conducting an evaluation is guided by principles of inclusion, dialogue, and deliberation. Implementation of these principles is intended to “extend impartiality by including relevant interests, values, and views so that conclusions can be unbiased in value as well as factual aspects”

(House and Howe, 2000, p. x). In the context of the guiding principles, the evaluator is responsible for systematic and unbiased data collection, analysis, and presentation of findings.

Gary T. Henry. The ultimate goal of evaluation, social betterment, drives the ERE approach proposed by Mark, Henry, and Julnes (1998, 2000). ERE focuses on understanding the underlying mechanisms of programs and thus is concerned primarily with addressing the research questions that help identify program mechanisms that are operating and those that are not (Mark, Henry, and Julnes, 1998). The ERE evaluator develops a study in an attempt to understand causality rather than being satisfied with a description of program outcomes. As a result, the ERE evaluation attempts not only to identify well-functioning program mechanisms, but also to enhance the generalizable knowledge base of a particular set of programs or program theories.

The emergent realist evaluator takes into account program effects that are of most interest to the public and other relevant stakeholders and thus must determine stakeholders' values when investigating possible mechanisms. (This part reflects the attention given to the values of stakeholders when conducting and framing evaluation questions.) Henry (1996) argues that understanding the values of the various stakeholder groups, including areas of consensus and conflict, helps to promote democratic policy. ERE offers three methods for investigating stakeholder values. The first involves surveying and sampling possible stakeholders, and the second uses qualitative methods such as interviews or focus groups (or both) to determine their needs and concerns. The third approach involves analyzing the context of the evaluation from a broad philosophical perspective, focusing on issues such as equity, equality, and freedom. (This third approach is referred to as "critical analysis.") These value investigations should then be communicated to the multiple audiences of the evaluation (Mark, Henry, and Julnes, 1998).

The ERE evaluator also engages in a process of *competitive elaboration* or *principled discovery*. Competitive elaboration involves ruling out alternative explanations for study findings, which includes exploring alternative program theories and threats to validity (Mark, Henry, and Julnes, 1998). This requires a preexisting body of knowledge of possible program mechanisms and, ideally, a study design that experimentally tests each relevant mechanism to identify those with the greatest or least impact. (Relevant mechanisms are determined by a combination of stakeholder involvement and existing theories about a program.) This approach lends itself to quantitative methods of inquiry (such as fully randomized experiments and quasi-experiments). Principled discovery is used when programs are evaluated before practitioners are able to develop experientially tested theories (Mark, Henry, and Julnes, 1998). Approaches to discovering program mechanisms include exploratory data analysis (Tukey, 1977), graphical methods (Henry, 1995), regression analysis, and a context-confirmatory

approach (Julnes, 1995). (Under the context-confirmatory approach, “an empirical discovery that suggests a mechanism is used to generate a distinct prediction that should be true if the newly induced mechanism is operating”; Henry, Mark, and Julnes, 1998, p. 14.)

Stewart I. Donaldson. Donaldson is a proponent of what he refers to as *program theory-driven evaluation science* (PT-DES). The phrase *program theory-driven* (instead of *theory driven*) is intended to clarify the meaning of the use of the word *theory* in the evaluation context of evaluation by attempting to specify the type of theory (for example, program theory, not necessarily social science theory) that is guiding the evaluation questions and design. Donaldson distinguishes between program theory and other types of theory, such as social science theory, and defines program theory as a theory of how program components are presumed to affect outcomes and the conditions under which these processes are believed to operate (Donaldson, 2001). *Evaluation science* (instead of evaluation) is intended to underscore the use of rigorous scientific methods (qualitative, quantitative, and mixed method designs) to attempt to answer valued evaluation questions. It also is intended to highlight the emphasis placed on the guiding principle of systematic inquiry (American Evaluation Association, 1995) and the critical evaluation standard of accuracy (Joint Committee on Standards for Educational Evaluation, 1994).

PT-DES is the systematic use of substantive knowledge about the phenomena under investigation and scientific methods to determine the merit, worth, and significance of evaluands such as social, educational, health, community, and organizational programs (Donaldson, forthcoming). Its application involves using program theory to define and prioritize evaluation questions and using scientific methods to answer those questions.

Donaldson offers a simple three-step model for understanding the basic activities of program theory-driven evaluation science:

1. Developing program theory
2. Formulating and prioritizing evaluation questions
3. Answering evaluation questions

Simply stated, evaluators typically work collaboratively with stakeholders to develop a common understanding of how a program is presumed to solve a problem of interest. Social science theory and prior research (if they exist) can be used to inform this discussion and assess the feasibility of the proposed relationships between a program and its desired initial, intermediate, and long-term outcomes (Donaldson and Lipsey, forthcoming). This common understanding of program theory helps evaluators and stakeholders identify and prioritize evaluation questions. Evaluation questions of most interest are then answered using the most rigorous scientific methods possible given the practical constraints of the evaluation context.

Donaldson suggests that one of the best examples to date of program theory-driven evaluation science in action is embodied in the Centers for Disease Control's (CDC) six-step Program Evaluation Framework (1999). He argues that this framework is not only conceptually well developed and instructive for evaluation practitioners, but also has been widely adopted for evaluating federally funded public health programs throughout the United States. The CDC framework begins to unpack the more concise three-step PT-DES model and offers more details to help guide practitioners:

1. *Engage stakeholders*—Those involved, those affected, primary intended users
2. *Describe the program*—Need, expected effects, activities, resources, stage, context, logic model
3. *Focus the evaluation design*—Purpose, users, uses, questions, methods, agreements
4. *Gather credible evidence*—Indicators, sources, quality, quantity, logistics
5. *Justify conclusions*—Standards, analysis and synthesis, interpretation, judgment, recommendations
6. *Ensure use and share lessons learned*—Design, preparation, feedback, follow-up, dissemination

The primary focus of PT-DES is on the development of program theory and evaluation questions. As a result, PT-DES is arguably method neutral. That is, the focus on program theory often creates a superordinate goal that helps evaluators move beyond the debates about which methods are superior to use in evaluation practice (Donaldson, forthcoming; Donaldson and Christie, 2005; Donaldson and Scriven, 2003). From this contingency point of view, it is believed that quantitative, qualitative, and mixed method designs are neither superior nor applicable in every evaluation context (Chen, 1997). Whether an evaluator uses case studies, observational methods, structured or unstructured interviews, online or telephone survey research, a quasi-experiment, or a randomized experimental trial to answer the key evaluation questions is dependent on discussions with relevant stakeholders about what would constitute credible evidence in this context and what is feasible given the practical and financial constraints.

Jean A. King. King is concerned with designing and implementing evaluations in a collaborative manner with stakeholders for the purpose of increasing the likelihood that the information generated from the evaluation will be used. This collaboration involves stakeholder participation throughout the evaluation process. Her theoretical approach, *interactive evaluation practice* (IEP), emphasizes participation, use, capacity building, and the “interpersonal factor” (King and Stevahn, 2005).

King describes the IEP framework as “the intentional act of engaging people in making decisions, taking action, and reflecting while conducting

an evaluation study" (King and Stevahn, 2005). At its core, IEP is a participatory process. King (1998) argues that for a participatory evaluation to succeed, several conditions must be met. There must be an accepting power structure; shared meaning of experiences among participants, volunteers, and leaders; a great degree of interpersonal and organizational trust; enough time; and enough resources. An accepting power structure, King explains, is one that is open and willing to engage in the evaluation process and to use evaluative information to improve programs. For the evaluation to be successful, King underscores the importance of communication and discussion among stakeholders in creating shared meaning. She maintains that this process promotes interest in formulating methods of inquiry and collecting and interpreting data. The IEP evaluator also needs to identify and foster leaders during the evaluation process. Leaders, King argues, are needed to attract or recruit people to the evaluation process. They should be eager to learn and facilitate the evaluation process and "willing to stay the course when things go wrong" (King, 1998, p. 64). According to King, trust is a necessary condition for a successful participatory evaluation, and she urges evaluators to pay close attention to the interpersonal dynamics occurring during evaluations. She describes an evaluation without effective interpersonal interaction as "a machine without proper lubrication" (King and Stevahn, 2005).

The IEP evaluator has three primary roles: decision maker, actor, and reflective practitioner. As a decision maker, the evaluator is compelled to make decisions to facilitate the progress of the evaluation. Thus, the evaluator herself does not make crucial decisions about the program (that is, formative changes); rather, she makes decisions that propel the evaluation forward. As an actor, King explains, the evaluator participates in the evaluation process—which she describes as a "theatrical event"—as a "performer." That is, evaluators must see themselves as just one, albeit an important one, of many players in the evaluation process. The role of the evaluator, then, as a principal player, is that of a leader, a manager, and a "wise counselor." With her experience and expertise, the evaluator, as a reflective practitioner, also helps guide and assess evaluation progress. These three evaluator roles emphasize and reinforce the importance of the "personal factor" (Patton, 1997) when conducting evaluations.

Most recently, King has become more attentive to building evaluation capacity through participation as a means for ensuring ongoing evaluation use (King, 2002). This is foreshadowed in the work referenced here. The rationale for capacity building is that adults learn best by constructing personal meaning from their practice and can learn well in settings where they value both the task to be completed—in this case, the work of the school—and their collegial relationships (King and Stevahn, 2002). In addition to grounding in utilization focused evaluation, the origins of evaluation capacity building include the multiple traditions of action research (King and Lonnquist, 1994) and participatory evaluation and the concepts

of organizational learning and professional community. In a grounded ECB framework, King and Volkov (2005) identify three areas of focus: (1) organizational context, including internal and external factors; (2) resources, including access to evaluation resources and sources of support for evaluation in the organization; and (3) structures, including an oversight mechanism, a formal ECB plan, core ECB principles in policies and procedures, infrastructure to support the evaluation process, purposeful socialization, and a peer learning structure.

Conclusion

In this chapter, we provide an argument for inviting Greene, Henry, Donaldson, and King to participate in an academic exercise that is intended to bring evaluation theories to life. We briefly summarize the main thrust of each of the theorists' models, with the intent of providing a general understanding of their approaches. This summary also serves as a frame for understanding the subsequent four chapters, in which each theorist describes how he or she would evaluate the Bunche–Da Vinci case presented in Chapter One.

What is exciting about the theorists' chapters is that each describes what they would actually do when designing an evaluation for this case. Evaluation theory literature largely addresses practice in the abstract. That is, theorists often argue what they believe to be the primary purpose of evaluation and how that purpose should shape the evaluation process. There is, of course, discussion of the principal components of the approach, but often these discussions are more conceptual. And this is understandable. There are many restrictions (such as page length) placed on authors when publishing manuscripts in journals and book chapters, the primary outlet for our theories. Thus, theorists are more or less mandated to discuss the defining features of their model, explaining what makes their approach unique. Consider, for example, House's earlier writings on social justice and evaluation. He spoke prolifically of inclusion and argued for increased participation of underrepresented groups in the evaluation process. He forced us to acknowledge that evaluation does in fact determine who gets what, and thus the need to represent all relevant stakeholders (most important, those who typically do not have a voice in our society) in the evaluation process equally. Yet it was difficult to imagine from these writings what House would actually do, step-by-step, when evaluating, for instance, a statewide education initiative.

In this chapter, we do not present a how-to description of theorists' approaches, in part because it is rare for theorists to publish such portrayals of their models. Instead, most theorists publish frameworks for action that describe the general approach the theorist espouses. This is because, by and large, evaluation designs and the implementation of these designs must be tailored to fit the specific program context. The particularities introduced

by program context makes a one-size-fits all—it can work anywhere at any time—approach to evaluation relatively suspect. Thus, it is arguably more appropriate to present practical theories as general frameworks for actions that are meant to be adapted by and to circumstance. The how-to description, however, is of great interest to evaluators, as it helps bring color to some of the seemingly gray elements of a general framework. With this volume, the editors hope to introduce a bit of color to the theoretical landscape by presenting descriptions, written by the theorists themselves, of how to use their models in a specific program context.

References

- Alkin, M. C., and Christie, C. A. "An Evaluation Theory Tree." In M. C. Alkin (ed.), *Evaluation Roots*. Thousand Oaks, Calif.: Sage, 2004.
- Alkin, M. C., and House, E. "Evaluation of Programs." In M. C. Alkin (ed.), *Encyclopedia of Educational Research*. (6th ed.) Old Tappan, N.J.: Macmillan, 1992.
- American Evaluation Association. "Guiding Principles for Evaluators." In W. R. Shadish, D. L. Newman, M. A. Scheirer, and C. Wye (eds.), *Guiding Principles for Evaluators*. New Directions for Program Evaluation, no. 34. San Francisco: Jossey-Bass, 1995.
- Chen, H. "Applying Mixed Methods Under the Framework of Theory Driven Evaluations." In J. C. Greene and V. J. Caracelli (eds.), *Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms*. New Directions for Evaluation, no. 74. San Francisco: Jossey-Bass, 1997.
- Chen, H. T. *Theory-Driven Evaluations*. Thousand Oaks, Calif.: Sage, 1990.
- Cousins, J. B., and Whitmore, E. "Framing Participatory Evaluation." In E. Whitmore (ed.), *Understanding and Practicing Participatory Evaluation: New Directions for Evaluation*. New Directions in Evaluation, no. 80. San Francisco: Jossey-Bass, 1998.
- Donaldson, S. I. "Overcoming Our Negative Reputation: Evaluation Becomes Known as a Helping Profession." *American Journal of Evaluation*, 2001, 22, 355–361.
- Donaldson, S. I. "Theory-Driven Program Evaluation in the New Millennium." In S. Donaldson and M. Scriven (eds.), *Evaluating Social Programs and Problems: Visions for the New Millennium*. Mahwah, N.J.: Erlbaum, 2003.
- Donaldson, S. I. *Program Theory-Driven Evaluation Science: Strategies and Applications*. Mahwah, N.J.: Erlbaum forthcoming.
- Donaldson, S. I., and Christie, C. A. "The 2004 Claremont Debate: Lipsey vs. Scriven. Determining Causality in Program Evaluation and Applied Research: Should Experimental Evidence Be the Gold Standard?" *Journal of Multidisciplinary Evaluation*, 2005.
- Donaldson, S. I., and Lipsey, M. W. "Roles for Theory in Evaluation Practice." In I. Shaw, J. Greene, and M. Mark (eds.), *Handbook of Evaluation*. Thousand Oaks, Calif.: Sage, forthcoming.
- Donaldson, S. I., and Scriven, M. "Diverse Visions for Evaluation in the New Millennium: Should We Integrate or Embrace Diversity?" In S. I. Donaldson and M. Scriven (eds.), *Evaluating Social Programs and Problems: Visions for the New Millennium*. Mahwah, N.J.: Erlbaum, 2003.
- Fetterman, D. M. *Foundations of Empowerment Evaluation*. Thousand Oaks, Calif.: Sage, 2001.
- Fitzpatrick, J. L., Sanders, J. R., and Worthen, B. R. *Program Evaluation: Alternative Approaches and Practical Guidelines*. (3rd ed.) Needham Heights, Mass.: Allyn and Bacon, 2003.
- Greene, J. G. "Challenges in Practicing Deliberative Democratic Evaluation." In R. R. Goetz, B. H. McFarland, and K. V. Ross (eds.), *What the Oregon Health Plan Can Teach*

- Us About Managed Mental Health Care*. New Directions for Mental Health Services, no. 85. San Francisco: Jossey-Bass, 2000.
- Greene, J. G., and Caracelli, V. J. "Defining and Describing the Paradigm Issues in Mixed-Method Evaluation." In J. C. Greene and V. J. Caracelli (eds.), *Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms*. New Directions for Evaluation, no. 74. San Francisco: Jossey-Bass, 1997.
- Henry, G. T. *Graphing Data: Techniques for Display and Analysis*. Thousand Oaks, Calif.: Sage, 1995.
- Henry, G. T. "Does the Public Have a Role in Evaluation? Surveys and Democratic Discourse." In M. Braverman and J. K. Slater (eds.), *Advances in Survey Research*. New Directions for Evaluation, no. 70. San Francisco: Jossey-Bass, 1996.
- House, E., and Howe, K. "Deliberative Democratic Evaluation." In R. R. Goetz, B. H. McFarland, and K. V. Ross (eds.), *What the Oregon Health Plan Can Teach Us About Managed Mental Health Care*. New Directions for Mental Health Services, no. 85. San Francisco: Jossey-Bass, 2000.
- Joint Committee on Standards for Educational Evaluation. *The Program Evaluation Standards*. Thousand Oaks, Calif.: Sage, 1994.
- Julnes, G. "Context-Confirmatory Methods for Supporting Disciplined Induction in Post-positivist Inquiry." Paper presented at the annual meeting of the American Evaluation Association, Vancouver, Canada, 1995.
- King, J. A. "Making Sense of Participatory Evaluation Practice." In E. Whitmore (ed.), *Understanding and Practicing Participatory Evaluation: New Directions for Evaluation*. New Directions for Evaluation, no. 80. San Francisco: Jossey-Bass, 1998.
- King, J. A. "Building Evaluation Capacity in a School District." In D. W. Compton, M. Baizerman, and S. H. Stockdill (eds.), *The Art, Craft, and Science of Evaluation Capacity Building*. New Directions for Evaluation, no. 93. San Francisco: Jossey-Bass, 2002.
- King, J. A., and Lonnquist, M. P. *Learning from the Literature: Fifty Years of Action Research*. Madison, Wis.: Center for the Organization and Restructuring of Schools, 1994.
- King, J. A., and Stevahn, L. "Three Frameworks for Considering Evaluator Role." In K. E. Ryan and T. A. Schwandt (eds.), *Exploring Evaluator Role and Identity*. Greenwich, Conn.: Information Age Publishing, 2002.
- King, J. A., and Stevahn, L. "Interactive Evaluation Practice: Managing the Interpersonal Dynamics of Program Evaluation." Unpublished manuscript, 2005.
- King, J. A., and Volkov, B. "A Grounded Framework for Evaluation Capacity Building." Unpublished manuscript, University of Minnesota, 2005.
- Mark, M. M., Henry, G. T., and Julnes, G. "A Realist Theory of Evaluation Practice." In G. T. Henry, G. Julnes, and M. M. Mark (eds.), *Realist Evaluation: An Emerging Theory in Support of Practice*. New Directions for Evaluation, no. 78. San Francisco: Jossey-Bass, 1998.
- Mark, M. M., Henry, G. T., and Julnes, G. *Evaluation: An Integrated Framework for Understanding, Guiding, and Improving Public and Non-Profit Policies and Programs*. San Francisco: Jossey-Bass, 2000.
- Patton, M. Q. *Utilization-Focused Evaluation: The New Century Text*. (3rd ed.) Thousand Oaks, Calif.: Sage, 1997.
- Preskill, H., and Torres, R. T. *Evaluative Inquiry for Learning in Organizations*. Thousand Oaks, Calif.: Sage, 1999.
- Rossi, P. H., and Freeman, H. E. *Evaluation: A Systematic Approach*. (5th ed.) Thousand Oaks, Calif.: Sage, 1993.
- Shadish, W. R., Cook, T. D., and Leviton, L. C. *Foundations of Program Evaluation: Theories of Practice*. Thousand Oaks, Calif.: Sage, 1991.
- Stake, R. E. "Program Evaluation, Particularly Responsive Evaluation." In W. B. Dockrell and D. Hamilton (eds.), *Rethinking Educational Research*. London: Hodder and Stoughton, 1980.

- Stake, R. E. "Program Evaluation, Particularly Responsive Evaluation." In G. F. Madus, M. S. Scriven, and D. L. Stufflebeam (eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*. Norwood, Mass.: Kluwer-Nijhoff, 1987.
- Stake, R. E. "Stake and Responsive Evaluation." In M. C. Alkin (ed.), *Evaluation Roots*. Thousand Oaks, Calif.: Sage, 2004.
- Tukey, J. W. *Exploratory Data Analysis*. Reading, Mass.: Addison-Wesley, 1977.
- Weiss, C. H. *Evaluation: Methods for Studying Programs and Policies*. (2nd ed.) Upper Saddle River, N.J.: Prentice Hall, 1998.

CHRISTINA A. CHRISTIE is an assistant professor and associate director of the Institute of Organizational and Program Evaluation Research at Claremont Graduate University.

TAREK AZZAM is a doctoral student in the Social Research Methods Division in the Graduate School of Education and Information Studies at the University of California, Los Angeles.