

Preface

FOCUS OF THE BOOK

There is an old adage that civilization must continually find new solutions to the persistent underlying problems of food, shelter, health, and communication. Similarly, in evaluation there are recurring themes, problems, and issues that periodically resurface in new forms to demand our attention.

- Why should evaluation be done? To improve programs? To influence decision making? To protect the public? To solve social problems? To promote social diversity?
- What are the proper social role(s) for the evaluator as a professional? As a researcher? As a teacher? As an advocate? As a facilitator? As a judge?
- How can stakeholders best be involved in evaluation studies? As served clients? As participants? As collaborators? As empowered citizens?
- What should we consider acceptable evidence for making evaluative decisions? Causal claims? Moral conclusions? Expert opinion? Aesthetic judgments? Stakeholder consensus?
- How do we arrive at the most valid understandings of quality? Controlled experiments? Moral deliberation? Phenomenological renderings?
- What is the most effective way to ensure the quality of evaluation practice? Advanced training? Accreditation and licensing? Consensual professional standards? Mandatory metaevaluation?

Fundamental issues are therefore those *underlying concerns, problems, or choices that continually resurface in different guises* throughout evaluation work. By their very nature, these issues can never be finally solved but only temporarily resolved. Fundamental issues underlie all areas of evaluation, whether it be communication with clients, ethical dilemmas, cultural differences, preparation of new evaluators, work with special populations, governmental service, methodological difficulties, social justice, evaluation influence, or economic survival as a professional. In this volume, we examine the nature and importance of issues such as these, issues that influence the character of evaluation theory, method, practice, and, indeed, the profession itself.

The chapter authors in this volume identify such issues in evaluation work, examine their importance, reflect on how they impact the nature of evaluation, and consider effective means of dealing with them. Such examinations help keep current problems in evaluation in better historical perspective, provoke thoughtful consideration of present options, and enable us to create more effective alternatives for the future. The considerations of fundamental issues in evaluation that are raised in this volume provide lively topics for analysis, disagreement, and discussion. They include issues that will continue to shape the nature and future of the evaluation profession.

DEVELOPMENT OF THE BOOK

“Fundamental Issues in Evaluation” was the conference theme of the 2004 annual meeting of the American Evaluation Association (AEA). This theme was the topic of plenary presentations made at the conference as well as the focus of a collection of Presidential Strand presentations featured throughout the meeting. Smith, as the 2004 AEA president, and Brandon, as the presidential strand chairperson, collaboratively developed the theme and invited a wide range of conference presentations on the topic. A few of the best of these presentations were solicited for inclusion in the present volume. Additional chapters were subsequently solicited to ensure a balanced coverage of the overall topic.

The initial conference presentations have been substantially reworked and expanded. Critical reviews of interim drafts were provided by the volume editors as well as by independent reviewers selected for their expertise in the respective topics of each chapter. These reviewers were selected mutually by the chapter authors and the volume editors. We are pleased to acknowledge below the important contributions of

these independent reviewers. The result of this careful attention to quality is this impressive collection of chapters examining a range of issues that have played, and will continue to play, a central role in the profession of evaluation.

ORGANIZATION OF THE BOOK

In order to understand better the different foci of fundamental issues in evaluation, we have identified four general overlapping categories: theory, method, practice, and profession. These categories were used in scheduling the AEA conference presentations and are included here as an organizing structure for the volume. These four aspects encompass the world of evaluation, but they are simultaneously separate and interconnected. Some fundamental issues may be primarily problems of theory or practice, but many are entangled across several or all of these aspects. A category description focuses attention on a dominant aspect of the fundamental issue, but fundamental issues are often enmeshed across two or more categories.

Theory is that aspect reflecting our thinking about how and why we engage in evaluation; whether evaluation is done for purposes of validation, accountability, monitoring, or improvement and development; whether evaluation is a form of knowledge production, client service, social reform, or political control.

Method reflects our tools, the means by which we accomplish our intended purposes; tools of inquiry, of resource management, of interpersonal engagement, of social reform.

Practice reminds us of the immediate world of politics, clients, resources, role ambiguity, and changing field conditions; the practical concerns of getting the work done well and of making a difference.

Profession is about evaluation as a socially defined guild of practitioners and colleagues; the definitions of our work, our ethics, our responsibilities, and our livelihoods.

Because fundamental issues are, well, fundamental, any specific issue tends to have implications across all four aspects. For example, whether to employ a particular strategy for involving stakeholders in an evaluation might raise considerations of (1) whether the strategy is, in principle, well designed (method), (2) whether it can be effectively implemented in the current situation (practice), (3) whether its use is consistent with what is considered to be the proper role of stakeholder involvement in this particular evaluation (theory), and (4) the extent to which its use might benefit or harm society (profession). Which aspect is con-

sidered most salient depends, of course, on how one frames the issue—a methodologist may see the issue as one of designing an effective strategy (method), while a social critic may see the issue as whether use of the strategy promotes social justice (profession).

The fundamental issues discussed in this volume could similarly be classified across several or all four aspects, depending on one's point of view. For example, some issues that arise as concerns of method, such as in Chapter 4, "Multiple Threats to the Validity of Randomized Studies," can be shown to have profound implications for the profession of evaluation itself, as seen in Chapter 10, "The Impact of Narrow Views of Scientific Rigor on Evaluation Practices for Underrepresented Groups."

Each of the chapters in this volume alludes to multiple fundamental issues, issues that reflect complex, interwoven concerns that evolve over time. The grouping of the chapters in this volume reflects in part, therefore, how the present authors are currently dealing with these fundamental concerns, and what we, as editors, see as important interconnections. The fundamental issues represented here, however, will likely resurface in the future with possibly some other aspect seen as more salient at that moment in time. Although we have placed the chapters into sections of theory, method, practice, and profession according to what we judge to be the dominant perspective in each chapter, we invite the reader to consider the implications of each chapter for the other aspects of evaluation.

After an opening chapter, the book is divided into four parts. Part introductions are included in the volume to introduce the chapters and to highlight how each contributes to understanding of the relevant aspect of theory, method, practice, or profession considered in that part. There are, of course, scores of fundamental issues possible under each aspect; those included in this volume are but a sample of some of the fundamental issues currently of most concern in the evaluation profession.

CONTENTS OF THE BOOK

In "Fundamental Issues in Evaluation" (Chapter 1), Nick L. Smith introduces the topic of fundamental issues in evaluation. He presents an overview, identifies a variety of examples, considers the characteristics of fundamental issues, and discusses why and how such issues are fundamental in evaluation. This chapter explains why the concept of fundamental issue is a productive way to examine recurring themes and prob-

lems in evaluation, and it provides a conceptual framework within which to consider subsequent chapters of the volume. Each chapter treats one or more fundamental issues, grouped according to the aspects of theory, method, practice, and profession.

Part I: Issues of Theory includes two chapters. In “The Relevance of Practical Knowledge Traditions to Evaluation Practice” (Chapter 2), Thomas A. Schwandt raises the question, What is the nature of practical knowledge and how does it influence our understanding and practice of evaluation? He argues that good practice in evaluation “depends in a significant way on the experiential, existential knowledge we speak of as perceptivity, insightfulness, and deliberative judgment” (p. 37). In considering the fundamental issue, What is the nature of expertise in evaluation?, Schwandt contrasts the role of practical knowledge with more technicist interpretations.

In her chapter, “Stakeholder Representation in Culturally Complex Communities: Insights from the Transformative Paradigm” (Chapter 3), Donna M. Mertens raises the difficult recurring issues of how to deal with cultural competence and the treatment of diversity within evaluation practice. She proposes the use of a transformative paradigm that seeks to ensure that stakeholders are fairly represented, that the diversity of stakeholder groups is considered adequately, and that differences in power among these groups are addressed in evaluations.

Both chapters in Part I illustrate how selecting a particular theoretical position about the nature of evaluation subsequently shapes the way we think about the profession itself and therefore which methods and forms of practice we consider appropriate.

“Multiple Threats to the Validity of Randomized Studies” (Chapter 4) by Judith A. Droitcour and Mary Grace Kovar leads off *Part II: Issues of Method*. The authors address the fundamental issue, How can experimental studies best be designed for practical use in the field?, by focusing on two threats to validity when using experiments: differential reactivity and biasing social interactions. Using numerous examples, they present a clear, succinct summary of the literature on these threats, with the valuable addition of suggesting means for prospective identification of the two forms of bias.

The second of the three chapters in Part II is “Research Synthesis: Toward Broad-Based Evidence” (Chapter 5). In this chapter, Susan N. Labin presents an overview of types of research syntheses including traditional literature reviews, qualitative reviews, and broad-based research syntheses; meta-analyses; evaluation syntheses and retrospective evaluation syntheses; and the synthesis method presented in the *Guide to Community Preventive Services*, as well as others. Her critical review of the

advantages, disadvantages, and selection criteria of these methods provides important insights about the fundamental issue, What is the proper way to synthesize evaluation and social science research findings in evaluation?

“Building a Better Evidence Base for Evaluation Theory: Beyond General Calls to a Framework of Types of Research on Evaluation” (Chapter 6), by Melvin M. Mark, completes Part II. Mark provides a taxonomy of types of studies for researching evaluation, including descriptive studies, classification studies, causal analyses, and values inquiries. The taxonomy serves as a guide for the design, classification, and interpretation of research on evaluation. As Mark points out, a stronger empirical understanding of prior evaluation work can contribute to better responses to questions about other fundamental issues, such as What are the best approaches and methods for a given evaluation?, How have those choices been manifested in previous studies?, and What have been the consequences of those choices?

All three chapters in Part II provide theoretical arguments concerning which methods will most improve evaluation practice. Collectively they portray the profession of evaluation as an empirically based, inquiry-focused enterprise, and they each make important contributions to technical issues of evaluation method.

Part III: Issues of Practice begins with “Complexities in Setting Program Standards in Collaborative Evaluation” (Chapter 7) by J. Bradley Cousins and Lyn M. Shulha. Cousins and Shulha address the fundamental issue, What level of performance must an evaluand achieve for its performance to be considered adequate or satisfactory? They discuss aspects of selecting bases for comparison, determining values, setting standards, and ensuring cultural sensitivity when stakeholders and evaluators work together to conduct formative collaborative evaluations. In such evaluations, power over evaluation decision making is shared equally among the evaluators, who have technical skills, and program personnel, whose understanding of the substance, organization, and daily operations of the program is far greater than that of the program’s evaluators. Such evaluations should be ideal for explicitly and carefully deciding how well the evaluated program has to perform.

Carlos C. Ayala and Paul R. Brandon continue the focus on formative evaluation in Chapter 8, “Building Evaluation Recommendations for Improvement: Insights from Student Formative Assessments.” They consider the case of K–12 student formative assessment that is conducted by the classroom teacher for the purpose of providing immediate feedback to students about how to improve their learning. Ayala and Brandon present a formal approach to assessment called the Assessment Pedagogy Model, which uses assessment as an integral part of teaching.

They describe the several components of the model and review implications of the model for formative program evaluation in general.

Both these chapters illustrate how difficulties in evaluation practice not only have resulted in subsequent improvements to that practice but also have led evaluators to reconsider issues of theory and method. These chapters illustrate the essential contribution of self-reflective practice in producing an effective and useful profession.

Part IV: Issues of the Profession consists of three chapters. In “What Is the Difference between Evaluation and Research—and Why Do We Care?” (Chapter 9), Sandra Mathison provides new insights on the perennial questions of how social research and evaluation differ: What is the overlap of the purposes and methods of these two forms of inquiry? What knowledge and skill sets are required of evaluators versus social science researchers? Mathison’s analysis of the depictions of the differences between these two forms of inquiry, such as the extent to which the findings of studies result in decisions (vs. generalizations), shows that the differences depend on the context and the purpose of particular studies. Her chapter will help readers move from the overly simple question, What is the difference between research and evaluation?, to a more sophisticated consideration of the conditions under which research and evaluation share both similarities and differences.

The two remaining chapters in Part IV address important fundamental issues that have received increasing attention in recent years. With greater appreciation for the diversity within society and the increasing globalization of evaluation, evaluators are now more explicitly dealing with such fundamental issues as What is the nature of a culturally competent evaluation? and What are the proper role and means for evaluation in contributing to social justice?

In Chapter 10, “The Impact of Narrow Views of Scientific Rigor on Evaluation Practices for Underrepresented Groups,” Elmima C. Johnson, Karen E. Kirkhart, Anna Marie Madison, Grayson B. Noley, and Guillermo Solano-Flores argue that rigorous studies, although internally valid, are often not sufficiently robust, because they do not arrive at conclusions that properly take into account program and evaluation context, particularly the contextual factors that affect communities of peoples of color. They argue that the primary purposes of program evaluations and assessments should be to ensure social justice, enhance fairness, and address inequities in access to power. They illustrate how robustness can be increased by focusing on three issues: (1) how definitions of validity determine the operationalization of rigor, (2) how language theory can inform the assessment of second-language learners, and (3) how evaluations can advance social justice by using

multiple measures of accountability. Their discussion speaks to fundamental issues of the proper purposes, roles, and uses of evaluation.

The third chapter in Part IV is “Improving the Practice of Evaluation through Indigenous Values and Methods: Decolonizing Evaluation Practice—Returning the Gaze from *Hawai’i* and *Aotearoa*” (Chapter 11), by Alice J. Kawakami, Kanani Aton, Fiona Cram, Morris K. Lai, and Laurie Porima. The authors make a strong case that many evaluations in Polynesian settings have adapted a “Western approach” that has ignored or disrespected the role of community elders as the arbiters of quality, has failed to acknowledge and incorporate indigenous spiritual values, and has disregarded cultural practices such as valuing lineage. They demonstrate that these evaluations typically have not benefited the communities in which the programs have occurred. They argue that indigenous peoples have a right to evaluation methods that are culturally relevant, not simply because the results will be more culturally and epistemologically valid, or because such methods will lead to greater stakeholder support of the resultant findings, but because indigenous peoples have a moral right to self-determination, a right to influence the choice of methods for evaluating the programs that affect their lives.

All three chapters in Part IV illustrate how our understanding of the nature of evaluation as a profession continues to evolve. The latter two chapters draw our attention to especially serious issues about what the generative principles should be that define who evaluators are, what they do, and how they operate—fundamental issues of the evaluation profession.

We offer here a note on how to read this volume. Chapter 1 provides an overview and general discussion of the nature and importance of fundamental issues in evaluation. It might be a useful starting point. Used as a text, one might read this volume sequentially, moving from theory through method, practice, and profession; this pattern reflects the sequence in which one typically learns about evaluation. Those readers with specific interests in a given aspect, say method or practice, might proceed immediately to those sections. Or, the volume can be used as a general resource, with the reader focusing just on the specific topics of interest. We encourage the reader to review each of the chapters in the volume. We have worked hard with each author to ensure that the material is accessible to all readers, not just to those with a prior background or interest in the specific topic. We believe that an important value of this book is that all readers can gain a new appreciation for the significance and long-term impact on evaluation of fundamental issues that lie outside their usual areas of interest.

ACKNOWLEDGMENTS

This volume reflects a novel way of examining recurring themes in evaluation. We are most grateful to the contributing authors for their patience and goodwill in helping us draw out the fundamental issues in their own work. To assist the authors, we collaboratively selected independent reviewers who provided insightful, constructive, and in some cases extensively detailed comments on earlier chapter drafts. This volume is much stronger because of their collective efforts, and we wish to thank them for their professional diligence and collegial support. Our grateful thanks to:

Gary Bond
Indiana University–Purdue University Indianapolis

Christina Christie
Claremont Graduate University

David Cordray
Vanderbilt University

Lois-ellin Datta
Datta Analysis

Lorna Earl
Ontario Institute for Studies in Education / University of Toronto

Leslie Goodyear
Education Development Center, Inc.

Robin Miller
Michigan State University

Hallie Preskill
Claremont Graduate University

Darlene Russ-Eft
Oregon State University

Mary Ann Scheirer
Scheirer Consulting

Elizabeth St. Pierre
University of Georgia

Veronica Thomas
Howard University

Elizabeth Whitmore
Carleton University

Susan Kistler, Executive Director of the American Evaluation Association (AEA), provided invaluable assistance in planning and conducting the 2004 AEA conference, during which early drafts of most of the chapters in this volume were presented. Her technical staff produced the excellent renderings of our concept of the mandala of overlapping nautilus shells. Without her outstanding organizational skills and professional support, this volume would not have been possible.

Finally, we offer our sincere thanks to C. Deborah Laughton of The Guilford Press for her unwavering support of this project. Her patience, thoughtful guidance, and good humor were key to our ultimately finishing this volume. Any remaining limitations are, of course, our own responsibility.

NICK L. SMITH
PAUL R. BRANDON