PREFACE

There are several evaluation books that present a survey of the various evaluation approaches, models, and theories; provide guidance for facilitating or managing the evaluation process; and/or explore some of the important theoretical, ethical, practical, and political issues in evaluation. These are an excellent source for giving the relative newcomer a sense of the diverse approaches and ideas within the evaluation community.

There are also numerous texts that guide readers through the process of using social science research methods to answer evaluation questions. The steps involved (with some variations) are typically clarifying evaluation questions, developing an appropriate research resign, identifying or developing measures, collecting data, and presenting findings.

Evaluation, however, is much more than "applied social science research." E-*valu*-ation, as the term implies, involves not only collecting descriptive information about a program, product, or other entity but also using something called "values" to (a) determine what information should be collected and (b) draw explicitly evaluative inferences from the data, that is, inferences that say something about the quality, value, or importance of something. Put another way, research can tell us "what's so," but only evaluation can tell us "so what."

Evaluation theorists for years have advised us to "take into account" all relevant values and to use them in the interpretation of data collected as part of an evaluation (House & Howe, 1999). But especially for the relatively new evaluator (even one who is knowledgeable and experienced in research methodology), there is not a lot of guidance about how this is done.

Scriven has made by far the greatest contributions to the development of a unique logic and methodology that is truly evaluation specific. He addresses head-on the issue of which values should be considered relevant in an evaluation and where they should be applied (e.g., Scriven, 1991).

What is "evaluation-specific logic and methodology"? It is a set of principles (logic) and procedures (methodology) that guides the evaluation team in the task of blending descriptive data with relevant values to draw explicitly evaluative conclusions. An explicitly evaluative conclusion is one that says how good, valuable, or important something is rather than just describing what it is like or what happened as a result of its implementation.

Evaluation-specific methodology is absolutely essential for answering truly evaluative questions such as whether a certain program, policy, or product is (a) just good enough to buy, fund, or support; (b) significantly better than that; (c) clearly better than the other two options we are considering (or might have considered); and/or (d) an excellent example of "best practice." In contrast, evaluation-specific methodology is *not* necessary for answering nonevaluative research questions, that is, those that are not directly concerned with quality, value, or importance.

WHO AND WHAT THIS BOOK IS FOR

This book is designed primarily for the many practitioners who find themselves thrown into evaluation roles without the benefit of formal evaluation training. Even those with a good grounding in applied research methods find that there is a whole lot more to putting together a good evaluation than meets the eye. For these people, and for students of evaluation (who struggle with many of the same issues and often plan to be practitioners as well), this is the straight-talking evaluation methodology book you have been dying to add to your bookshelf.

The purpose of this book is to provide a "nuts-and-bolts" guide that covers some of the practical and methodological basics of doing an evaluation. It is designed to lead the evaluation team (which may include people internal to the organization, external contractors, and/or other key stakeholders) through the steps involved in doing a good evaluation.

The focus is on evaluation-specific logic and methodology (which address the issues described previously) rather than on what is already covered well in social science research texts (e.g., the development of measures and instruments for collecting data). From "Where do I start?" to "How do I pull all of this information together into a report for my client?" this book provides a step-by-step guide, including checklists, rubrics, and rules of thumb for doing a real evaluation.

Topics covered in this book include the following:

- How to identify the right criteria for your evaluation (using needs assessment and other strategies)
- How to figure out which criteria are more important than others
- How to blend a mix of qualitative and quantitative data with relevant standards and values to draw explicitly evaluative conclusions (i.e., say something specific about quality or value)
- How to pull together information about all of the strengths and weaknesses to answer the fundamental evaluation questions for the client
- How to evaluate an evaluation

WHAT IS NEW OR FRESH ABOUT THIS APPROACH?

Although the work presented in this book does seek to bring some fresh insights and ideas, it stands on the shoulders of some veritable giants in the field to whom I owe a debt of gratitude. In particular, the pioneering work of Scriven on the conceptualization of evaluation's unique logic and methodology has provided a solid theoretical foundation for what I have attempted here. No one has thought more deeply or critically about what evaluation really is, how it is different from other undertakings, and what very specific procedures are required to answer truly evaluative questions than has Scriven.

The work of both Carol Weiss and Michael Quinn Patton on the utilization of evaluation have also been highly influential in the development of the ideas in this book. The same is true of the contributions from the various authors and practitioners who use program theory in evaluation (particularly the innovative work of those who have critiqued and improved on theory-based evaluation, such as Patricia Rogers and Carol Weiss). I have also drawn quite heavily on my knowledge of personnel evaluation and other developments from industrial and organizational psychology, organization development, and organizational learning. Finally, the practical know-how gleaned from working on real-world projects alongside colleagues at The Evaluation Center (especially Daniel Stufflebeam, James Sanders, and Gary Miron) has been invaluable for the development of new ideas, as has the experience of grappling with applied evaluation problems myself.

What I have attempted here is to combine a number of ideas from these supposedly diametrically opposed theoretical "camps" in the evaluation field

and to augment them with some of the tools, techniques, and ideas I have developed myself and with colleagues. My goal was to spell out a coherent approach that draws on the best of the best, adds its own twists to fill the gaps, and includes practical nuts-and-bolts guidelines and tools that could be applied in a real-world setting.

Although Scriven's work is widely respected in the field and many of the concepts he has introduced are in common use (e.g., formative evaluation, summative evaluation, goal-free evaluation, meta-evaluation), it has always puzzled me that far fewer people seem to apply evaluation-specific logic and methodology in practice. One possible explanation is that the depth and sophistication of Scriven's writing in this area make it a challenging read for the relative newcomer—although an incredibly valuable one for those who make the effort. But even for those who do, many people still find it hard to *visualize* what evaluation logic and methodology would really "look like" in practice. One of my goals in this book is to help fill this gap, albeit with my own interpretations of how best to apply the concepts.

A second major barrier to getting evaluation logic and methodology more widely integrated into evaluation practice, in my view, has been the lack of perceived linkages to the various other evaluation approaches that have contributed so much to the field. For example, there has been a persistent misconception that these fundamental principles and procedures apply only to independent evaluations (i.e., evaluations that do not include organizational members on the evaluation team). The methodologies explained in this book are designed to be just as useful as a guide for a facilitated participatory or collaborative evaluation as for a fully external noninteractive evaluation (i.e., one that does not include stakeholders on the evaluation team). Although many writers tend to promote the use of one approach or the other, I take the view that both approaches are valuable, depending on what primary need the evaluation is addressing (e.g., building organizational learning capacity, meeting accountability requirements) and on the time and resources available to get the evaluation done.

Another common (and completely incorrect) view is that the use of evaluation logic and methodology is somehow the antithesis of **theory-based evaluation** (i.e., evaluation that uses as its guiding framework the mechanism by which the program is expected to achieve its effects). In fact, this is one of the more powerful blends possible and is one of the key avenues I have pursued in an effort to develop some of the new evaluation methodologies that are presented in this book.

THE SIMPLICITY–COMPLEXITY TRADE-OFF

The intended approach of this book is to present the methodology that is specific to evaluation in as simple and straightforward a way as possible so that readers can see clearly how it works. There are, of course, many circumstances in which the evaluation waters are a lot muddier than might appear to be implied in much of this book. In particular, the higher up into the policymaking stratosphere we go (especially with public policy), the more we are dealing with difficult trade-offs among conflicting stakeholder values that cannot easily be resolved with, say, a high-quality needs assessment. Rather than confuse the newcomer with the nested cans of worms one might pry open here, I have (for the most part) deliberately chosen to set these issues aside while we first get on with the important task of understanding the basics of evaluation methodology.

This book is not intended to be a replacement for the more in-depth theoretical contributions already made by Scriven, Weiss, Patton, Rogers, Stufflebeam, or any of the others whose work I have drawn on. At the end of each chapter, a set of additional readings is listed for readers who wish to delve deeper into the concepts presented.

WHAT IS "THE TRUTH" IN EVALUATION?

One of the major points of departure among evaluators as they approach their work is grounded in how they view the notion of "truth" in evaluation. I take a fairly practical position on this that goes roughly as follows. Evaluation findings are "demonstrably true" when a solid mix of evidence supports a conclusion at or above the level of certainty required in that decision-making context. Does that mean that I think something has been "proven" to a 100% level of certainty so that it can be called an absolute truth in the strictest sense of the word? Certainly not. Every decision-making context has somewhat different requirements for the level of certainty needed for evaluation findings (similar to the standards of proof used to determine criminal responsibility vs. civil responsibility). In some cases, the stakes are very high and we need to be very sure indeed about our findings, at least "beyond reasonable doubt." In other cases, such a high level of certainty is not required and a "balance of evidence" standard of proof would be more appropriate. Most of my evaluation work to date has required levels of certainty somewhere between those two levels.

Those readers who have had a smattering or more of philosophy will recognize this as a *postpositive perspective.* In this book, quality or value is treated as a real (i.e., factual) attribute of things within a certain context, but one that is often hard to pin down. Getting "certain enough" answers to evaluation questions involves some serious detective work, including using multiple methods to unearth and synthesize the multiple perspectives that together can get us a close enough approximation to the truth.

The principle of obtaining multiple perspectives to get close enough to the truth applies not just to data collection but also to the all-important "values" side of evaluation. This issue is discussed in some detail in Chapter 6. Although some people might advocate addressing the values question in an opening chapter, based on my experience in teaching evaluation to relative newcomers to the profession, I am assuming that most people will want to get into the nuts and bolts before they tackle any tricky philosophical issues. However, those who find themselves distracted by this burning unanswered question should skip ahead to Chapter 6 early on and then come back to the nuts and bolts.

A MULTIDISCIPLINARY APPROACH

The purpose of this book is to contribute to the development of the "core" of evaluation—that which makes evaluation unique and that applies across multiple contexts, fields, and types of evaluands. To maximize relevance for a variety of applications, I have included diverse examples from program, policy, and project evaluation; personnel selection; performance appraisal; and product and service evaluation. This is not to say that each of these receives equal treatment in the book; the examples are much more skewed toward the evaluation of programs to reflect the interests of the largest audience.

ACKNOWLEDGMENTS