Being an EVALUATOR

Your Practical Guide to Evaluation

Donna R. Podems

THE GUILFORD PRESS
New York London

A Division of Guilford Publications, Inc.
370 Seventh Avenue, Suite 1200, New York, NY 10001
www.guilford.com

Printed in the United States of America

This book is printed on acid-free paper.

Last digit is print number: 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Names: Podems, Donna, author.
Title: Being an evaluator : your practical guide to evaluation / Donna R. Podems.
Description: New York : Guilford Press, [2019] | Includes bibliographical references and index.
Identifiers: LCCN 2018021202| ISBN 9781462537808 (pbk.) | ISBN 9781462537815 (hardcover)
Subjects: LCSH: Evaluation research (Social action programs) | Evaluation.
Classification: LCC H62 .P5725 2019 | DDC 001.4—dc23
LC record available at *https://lccn.loc.gov/2018021202*

The comics that appear in Chapters 3 and 6 and the icons were created by Christopher Lysy. Images on the game pieces are from Dreamstime.com, LLC.

I dedicate this book to four very important people in my life:

To my mom, Elizabeth Podems, who, since the first time
I could pick up a pencil, always encouraged me to write

To my dad, Gary Podems, who gave me the courage
to write what I wanted with that pencil

To my children, Gemma and Rhys, who bring amazing joy,
laughter, and fulfillment to my life— and without whom
I would have completed this book 3 years ago!

A Conversation with My Readers

I have written this book just for you: the practicing evaluator; the emerging evaluator; the "I am thinking about being an evaluator" nonevaluator; the "I do monitoring" evaluator; the "I was unceremoniously dumped into evaluation" evaluator; the "I just got out of grad school, and now I have to practice evaluation" evaluator; the "I do not do evaluations [yet]; I just need to know about evaluative thinking" nonevaluator; the researcher; and the person who needs to know how to work with evaluators.

I take each of you on a journey that demystifies evaluation, explores what it means to be an evaluator, and shares some well-kept trade secrets along the way. This is not an evaluation theory book, nor is it a research book. It is a book about being an evaluator who uses evaluation to explore, describe, explain, and eventually judge (in some way) how, and the extent to which, something does or does not work where, for whom, and why. The chapters offer various ways to engage with evaluation information and knowledge—ways that cement learning and encourage reflection. Structured guidance shows you how to untangle various evaluation situations through facilitation, negotiation, and listening. Some core aims of this book are to encourage you to experiment with different ways of thinking; consider multiple perspectives; and acknowledge and engage with the formidable roles that context, power, politics, culture, language, and values play throughout the *entire* evaluative journey.

How the Book Is Organized

I start at the beginning, as one should when writing a book, and then build each chapter on the previous one, providing scaffolding on which to engage in any evaluative process. However, the book does not need to be read sequentially. I have written the book to support your unique evaluation adventure, which may begin in Chapter 3 or Chapter 14, then jump to the last section of Chapter 5, and then circle back to a paragraph in Chapter 2. The book provides a structure that untangles the messiness of evaluation by informing and guiding the many choices faced, and the various discussions held, in an evaluation journey; at the same time, it allows for a realistic sense of unpredictability and differences found in each one.

The book has two parts. Part I lays the foundation for exploring any evaluative process: It breaks down and takes you through what an evaluator needs to know, whether it is an evaluative process for learning, reflection, improvement, accountability, social justice, or judgment. Part II focuses on working as an evaluator and exploring evaluation: It is aimed at fostering reflection on, and thoughtfulness and awareness of, the many kinds of evaluation that exist; discussing the many evaluative roles that can be filled; and exposing some common challenges and pitfalls often encountered in the field, but rarely described in textbooks. Fused together, Parts I and II demystify evaluation and provide a firm basis for candidly engaging with any monitoring or evaluation process, no matter what role you fulfill.

Throughout the book, I dive into the murky sea of evaluation and guide you through it all. Although I provide strategies and processes, I do not provide a "do this and then do that" model, which may work in some situations, and then not in others. Rather, I offer a way to *think* through any evaluative process so that you can comfortably engage with a peer, colleague, boss, beneficiary, or client in any evaluation situation. The book will support your work in almost any evaluation context—whether you work for a nonprofit, a community-based organization, a donor, a government, a university, an institute, or a foundation, or for some other group that aims to fix, change, influence, or in some way make something better in the social world.

THE LEARNING APPROACH

People learn in different ways and at different speeds. A concept that seems easy to grasp for you may be a stumbling block for others. This is true in most of life. Some people find it harder than others to tell time by using an analog clock, or instantaneously to know their left from their right. Often there is something in our daily lives that we find more challenging to do or understand than the people around us; often we do not talk about this. The same is true in evaluation. With this recognition, I draw on different ways of learning, including facilitated interactive activities, self-learning exercises, areas for reflection, sections for discussion, and practical applications. Furthermore, we all have lives outside of practicing evaluation. Some days I find I have a few minutes where I can quickly watch or listen to something, and some evenings I find an hour or two to read. Once in a blue moon, I find more focused time. Acknowledging that reality, I provide further ways to learn about each chapter's topic that meet your varying needs—carefully balancing how much more you want to learn with how much time you have to learn more.

Remember Aesop's fable about the Tortoise and the Hare? The overconfident, sleek, fancy-looking Hare bragged about how fast he could run. Tired of hearing him boast, the Tortoise challenged him to a race. The race was long and challenging, and the Hare sprinted ahead, while the Tortoise approached the race in a precise and methodical way. All the animals in the forest gathered to watch the race, which was won by the wise Tortoise. In the book, we (you and I) are Tortoises. We do not rush to the end to assess and judge; rather, we thoughtfully engage and delight in each step of the evaluative process.

My Perspective

Making sense of things in the social world is not an easy task for anyone. Being asked to judge something, and in doing so to value that thing, can be daunting. There is no perfect template. An evaluator needs to engage in the academic theory and the practical side of evaluation; it is not an either–or decision. Theory informs practice, and practice informs theory. An evaluator educates when appropriate, to ensure engagement, learning, and meaning in the process and in its results. An evaluator comfortably negotiates where necessary, and acknowledges and engages with culture, language, power, politics, and values that constantly swirl in the evaluation process and influence all decisions, including her own. (See "My Use of Pronouns," below.) Through demystifying the process, an evaluator invites people into it and enables them to join in; he actively takes that responsibility, and in doing so through evaluation makes the world a bit better off, even if just a tiny bit.

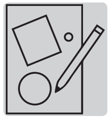
Why I Wrote the Book

The reason I am an evaluator is not that I enjoy judging things or being critical. I am an evaluator because I believe that evaluators have the potential to make the world a better place to live in. I have written this book to provide a way of thinking that supports an evaluator—not to intimidate others with knowledge, but to engage and guide them through a transparent process.

All over the world, especially where funds dedicated to social improvement are finite, thoughtful, kind, knowledgeable, and skillful evaluators can offer appropriate evaluative processes that inform decisions that affect people's lives, animals' lives, and the environment. Capable evaluators can fulfill an important societal role, while incompetent ones can be a detriment.

Some Guiding Icons

In an evaluation report, it is very common to have a page that lists all the abbreviations used in the report, along with their full terms; this provides a quick guide that you can refer to when reading the report. In this book, I use eight icons to draw your attention to how various features support you in more easily engaging with the text. The icon list follows.


Activities or things to do


Advice


Sharing a trade secret


Suggested resources


Clarifying questions


Practical conversations

THE USE OF PRONOUNS

The use of pronouns is a funny thing. In older books, there is a constant use of *he*. Being a *she* myself, I find that a bit bothersome. On the other hand, when I am reading, it is not always a smooth experience to read *he/she*. For some reason, when I read *he/she*, my brain always reacts with "Wait, so which is it? A he or a she?" This makes me stop reading the sentence, decide if it is a he or she, and then reread the sentence with my chosen pronoun. So I decided to alternate between using *he* and *she* throughout the book. But this then got me thinking: "What about people who do not identify with those pronouns? Perhaps I should use the pronoun *they*?" However, the use of *they* as a singular pronoun is not yet that common, and some early reviewers of the book kept thinking I was incorrectly using the plural *they*. So a note to my readers is that I use *he* and *she* alternately throughout the book, sprinkled with a few uses of *they*, to refer to a person.

Furthermore, I often switch in the book to using the term *we* instead of *you*, when you are technically not part of the *we* conversation just yet. I often find that using the term *you* can be heard as too direct, too confrontational, or too "othering." That is, it can be perceived as an attack that puts people on the defensive or makes them feel isolated. Of course, there are times when only *you* is appropriate, or indeed is the only option. However, using the term *we* underscores that we (you and I) are on the evaluation journey together.

Acknowledgments

If someone asked me to describe how this book came about, it would be a very long story. The story starts with my parents, Elizabeth and Gary Podems, who stood behind me and supported me in whatever I decided—whether it was passing up law school to join the Peace Corps, leaving a paying job to take a volunteer fellowship in South Africa, or living in Bosnia or Somaliland, because, as I explained to them, it just seemed right to me. All those journeys, and many more, informed the kind of evaluator I am. Mom inspired my love of writing by giving me my first journal at the age of 9, and then a new one every year after that.

Michael Quinn Patton entered my life at the beginning of my evaluation journey as my teacher, and has continued in the role of mentor and friend. Throughout my career, he has encouraged me not to conform to evaluation norms when they did not resonate with me (or my clients), to stay true to my principles and values, to march to the beat of my own drum (not his or anyone else's), and to write. He gave me the greatest gift: He believed in me.

I would not even be an evaluator had it not been for Paula Bilinsky, who plucked me out of nowhere and put her trust in a newbie just back from the Peace Corps; she patiently explained the basic ropes of evaluation, and has remained a constant and valued friend in my evaluation journey. Jennifer Greene grounded me in qualitative research and introduced me to feminist and democratic evaluation, all of which heavily influence my practice. She then went on to encourage me to edit my first book, which gave me the courage (and experience) to finish this one. After grad school, I approached C. Deborah Laughton and asked her guidance on how to publish a book; she provided a path that I diligently followed over the next 13 years, which proved to be excellent advice. I thank her for continued support during the making of this book.

Over the years, several of my colleagues and clients provided insights found in this book, as well as much-needed encouragement, for which I am most grateful. The book was a whisper of an idea when Catie Lott encouraged me. As it was taking shape, Kerstin Rausch Waddell provided a calming and insightful voice, which I relied on heavily throughout the journey. Anna Davis never stopped asking me when I was going to write the book, read multiple drafts, and gave me sage advice and thoughtful feedback. I stalled in the middle, and Tessie Catsambas provided a timely nudge. Benita Williams offered courage when I needed it most, along with her quantitative insights and the best-ever pep talks. Susan Tucker provided inspiring weekend and evening conversations that kept me on my toes. A book does not get written overnight (at least this one did not), and during the 5 years of writing, many other people have informed its shape and its content. I thank Leanne Adams, Daleen Botha, Cindy Clap-Wincek, E. Jane Davidson, Svetlana Negroustoueva, Tim Reilly, Gäelle Simon, Liezel de Waal, and Lauren Wildschut.

Robin Miller asked me to fill in for her and teach her evaluation foundation course at Michigan State University when I was a few years into writing the book. Teaching that class provided new insights, and Robin's thoughtful feedback on the second part of the book inspired me to think harder. Aimee N. White was supposed to be a "blind" peer reviewer; however, she knows my writing and stories so well that she guessed it was me. Her funny and sincere support, and meticulous feedback in the final, crucial stretch to the book's completion, were invaluable. I also thank the many peer reviewers who did not recognize me, and whose thoughtful criticism made this a better book: Mary E. Arnold, College of Public Health and Human Sciences, Oregon State University; Penny Burge, School of Education, Virginia Tech; Janice Fournillier, Department of Educational Policy, Georgia State University; John Klemanski, Department of Political Science, Oakland University; and Neil J. Vincent, Department of Social Work, DePaul University. Among the important elements of my book are its icons, comics, and game pieces, and I thank Christopher Lysy for enhancing the book with these fun and "totally me" additions.

Two very important people during the whole process were my children, Gemma and Rhys. I thank them for giving me the peace and quiet time to write (OK, I admit, it was often in exchange for iPad time) and for always asking, "How's the writing going, Mom?" Thank you, my Munchkins.

Contents

Purchasers can access an online supplement at *www.guilford.com/podems-materials* that contains some of the activities and games described in the book.

PART I
DOING EVALUATION AND THINKING EVALUATIVELY

An evaluator should understand how to gather data and use that to make a finding, a value judgment, and then develop a recommendation. I think being able to sustain that type of clearly articulated thinking throughout the conceptual pathway is a rare talent. . . .

—DUGAN FRASER, Chairperson of the South African M&E Association (2015)

CHAPTER 1
Speaking the Language

An evaluator is a human being living in a complex world. She (or he or they) engages with other human beings through a process that values something that is important to someone in some way. That process is called the **evaluative process,** and its findings are used to learn, improve, judge, or in some way inform decisions. How an evaluator thinks, who and what shapes her thinking, and how she views evaluation's role in society are all germane to how she defines herself as an evaluator. (For my use and choice of pronouns, please see page x in "A Conversation with My Readers" at the start of the book.)

Consider that the multidisciplinary field of evaluation includes pluralist philosophies that draw on multiple methods, criteria, measures, assessments, perspectives, audiences, and interests, as well as a plethora of traditions that have emerged from different social science disciplines (e.g., anthropology, psychology, sociology). Add to all this the roles of politics, power, culture, language, and values, and you can begin to see the many difficulties of definitively explaining what an evaluation is and how to design and implement one. Nonetheless, definitions and guidance abound, although of course they do not always agree.

Mix into this conversation the many roles that an evaluator can play, and you begin to see the complications of defining who is an evaluator, or drawing a boundary around what an evaluator does. Since the late 1980s, evaluation associations and societies, governments, international aid organizations, foundations, institutes, and other groups and institutions have been exploring ways to define who is an evaluator. However, evaluation differs from the medical, accounting, or legal fields, for example, in that no formal certification process is needed to practice evaluation: Anyone can claim to be an evaluator (McKegg, 2017; Podems & King, 2014).

Defining and clarifying what an evaluator does, how to be an evaluator, and the essence of being one constitute this book's core focus. Being comfortable with discussing the common language of evaluation provides a starting point from which to demystify evaluation and explore how to be an evaluator. If you would like to skip ahead to an in-depth discussion of what it means to be an evaluator, read Chapters 2 and 11. Otherwise, join me at the start of our evaluation journey, with a discussion of evaluation language.

The Power of Words

I begin with a memory of an awkward moment.

> "I grew up in a small town in New Jersey, and moved to South Africa in 1998. Back in those early days, I did not know too many people in South Africa. As the December holidays drew near, I received several invitations to holiday parties, and was especially glad to accept an invitation to join my neighbors for Christmas dinner. I was raised never to go to someone's house empty-handed, so I immediately asked what I could bring to the meal. My neighbor asked me to bring some crackers. I was glad for such an easy request. So I showed up on Christmas Day with two boxes of crackers; one salted and one flavored with rosemary. My hostess graciously accepted them. As we sat down to dinner, her husband loudly asked, 'Where are the crackers?' Since they were on the table, I pointed to them, at which point everyone burst out laughing. Apparently crackers are also small, prettily wrapped pieces of round cardboard, each containing a small prize; when you pull one open, it makes a small pop, like a firecracker. Who would have thought? I mean, who stops to ask, 'Tell me, what do you mean by a cracker?' "

Words may have different meanings to different people, in different contexts, and points in time. Sometimes the same word simply has multiple meanings for the same person. How and why people choose the words they use, and how their meanings are interpreted, are all part of everyday life; they are also relevant to any evaluative process. A critical part of being an evaluator is speaking the language. To do this, an evaluator needs to know the difference between evaluation words, terms, and concepts with concrete definitions that an evaluator should confidently know and be able to explain, such as *random sampling* (see Chapter 5), and words with meanings that vary by user and context, where evaluators often need to negotiate or clarify definitions and practical applications, such as the term *impact.* Although throughout the book I provide concrete definitions for evaluation terms and concepts that have them, the present chapter specifically looks at words and concepts with meanings and practical applications that need to be clarified prior to engaging in any evaluative process.

Language is a powerful tool, and a separate, broader conversation (which builds on the one found in this chapter) on how language influences evaluation is provided in Chapter 13. We start the evaluative journey here by exploring words, terms, and concepts that often perplex through their various definitions, interpretations, and applied meanings, and that have the potential to derail any evaluative process if they are not properly considered.

Before delving into examples for the evaluation field, let's start with some from everyday life. What would you think if someone said that the movie they just watched was "bad" or "sick"? Depending on the tone, the expression, the age of the speaker, and the cultural context, it would likely have different meanings, from "The movie was awful" to "It was fabulous." And because words have these nuances, there are always times in our lives where we feel we are the only ones in the room (or the world) who do not understand a term, or completely misunderstand it. Recently a younger colleague of mine asked what I was doing for the weekend. I responded that I was working on this book. She muttered something that sounded like "FOMO." I laughed and she laughed, but I was thinking, "What in the world is FOMO?" For those who do not know, it means <u>F</u>ear <u>O</u>f <u>M</u>issing <u>O</u>ut. I now know this, but I didn't know it then.

Consider a time when someone used a word or mentioned a concept that you thought you should know but did not, or you thought everyone else in the room knew and you did not. Most likely you nodded and pretended that you understood, and then did a web search on it later (or perhaps in the moment, if you had a chance to do so surreptitiously). Nearly everyone has been in this situation when it comes to working in the evaluation field. Ironically, it almost always starts with the mere mention of the field's most fundamental concept: evaluation.

The term **evaluation**; its cousin, **monitoring**; and the phrase and acronym that link them together, **monitoring and evaluation (M&E),** are among some of the more problematic terms in the field. While some definitions differ only slightly, others blatantly disregard and sometimes flat-out disagree with each other. These factors alone can create a breeding ground for intimidation (and perhaps insecurity, frustration, exasperation, or confusion, to name a few descriptors mentioned by colleagues and students) in the evaluation field. As one of my peers, Anna Davis, once told me, "In so many situations, I am sure I know what I am talking about with regard to monitoring or evaluation. Then someone comes along who has more power than me, uses a term in a different way, and tells me otherwise. . . . It is bullying by M&E!"

Thus an evaluator, or any person in a position of power, can indeed be a bully (either knowingly or unknowingly) or can be perceived as one when flinging about evaluation concepts or terms. Or the evaluator can choose to engage in a process that leads to a common understanding of a term. This approach lays the foundation for an evaluation process that is likely to be more useful to all involved.

BOX 1.1. Evaluative Process

Throughout Part 1, the term **evaluative process** is used to describe anything in an evaluation, from conceptualizing the intervention through to the judgment of it. Essentially, the evaluative process can be any part of a process where questions can be unearthed or answers can be found to address the "what," "when," "why," "where," "how," and "for whom" questions.

Why Evaluation Terms and Concepts Can Be Perplexing

Stand a little bit closer to me; I am going to share a trade secret. Evaluation terms and concepts abound, and for many, so do their definitions and their applications. More challenging is that sometimes differences in definitions are not always clear; the differences may lie in the nuances or the unspoken interpretation (as demonstrated through its application) of a definition in a specific setting. There are a few commonplace reasons for why this happens.

- Evaluation has emerged from diverse social science disciplines (e.g., anthropology, psychology, sociology), and these disciplines themselves use words and terms differently.
- Groups or individuals in positions of power, such as governments, donors, institutes, and foundations, are emboldened to define evaluation terms and concepts, and do so within their own worldviews. Their power stems from various sources; the two I mention here, money and influence, are tightly connected. Briefly, those who hold the purse strings can often define and apply words in any way they wish. If a program or intervention wants to receive funds, for example, more than likely it is the donor's definition that is used. This is often true even if that definition varies from other, more common understandings and uses.
- The innumerable evaluation textbooks, websites, evaluation guides, and blogs, which offer varying definitions of the same term, do not escape blame.

Why an Evaluator Should Clarify Terminology

Engaging explicitly with others about the meanings of evaluation terms and concepts *before* the start of any evaluative task will lay the foundation needed for a useful evaluation. These conversations help to avert misunderstandings, frustrations, potential conflicts, and possibly grave disappointment at the end. Furthermore, clarifying the terminology at the start of an evaluation process is a neutralizing activity. The start of an evaluative process is often nerve-wracking for those being evaluated, and creating the space to clarify terminology provides a neutral, nonjudgmental place in which to calm nerves and encourage dialogue and transparency.

Never assume that people understand evaluation terms or concepts, what they look like in practice, or how they influence an evaluation. When you are clarifying evaluation terminology, separate the terms into two groupings (Patton, 2012). The first group should contain words and concepts with concrete (though often misunderstood) definitions. Address these first. The first step for each of these terms is to provide a clear definition. The second step, and where the discussion is concentrated, is to describe how that term will be applied and how that application will influence the evaluation. An example of a word with a specific definition that commonly needs a thorough discussion is the term *causality*. The term has a concrete definition; there is no negotiation along the lines of "I think it means this, and he thinks it means that"; it means what it means. How it is applied in the evaluation is where the beneficial discussion happens. Do not assume that knowing the concrete definition automatically leads to an easy conversation; the conversation may still bring a challenging interaction. To read a technical discussion on causality, please skip ahead to Chapter 15.

The remainder of this chapter focuses on the second group: the words and concepts that have fuzzy and varying definitions. In each of these instances, discussing and negotiating the definition are wrapped up in the discussion of how it will be applied, making for an often complicated yet much-needed conversation. Regardless of the group into which a word falls, each conversation will, at the very least, support the evaluation's transparency, credibility, and use; *at the very most,* it will contribute to an engaging, collaborative, and social equity–building process.

The Evaluator's Role in Clarifying Terminology

When terms have multiple definitions and/or meanings, evaluators can often find themselves in the position of needing to clarify the terminology. As such, an evaluator can serve in three roles: as an educator, negotiator, *and* facilitator, so that each concept and its application are agreed upon for that specific evaluation. An intimate knowledge of the multiple definitions for the same term, and the subtler nuances that exist, will allow the evaluator to initiate and facilitate fruitful discussions. In turn, these discussions support equitable engagement, transparency, and clear expectations among those involved in the evaluative process of how the term will be applied, and how that will influence the evaluation. As suggested above, this process ideally increases the likelihood of a useful evaluation.

Throughout this book, the evaluator is positioned as one who leads processes to untangle meanings, clarify concepts, and concretize the application of evaluation terms, which supports evaluation use. In this way, I unashamedly lay the responsibility of a clear evaluation process in the hands of the evaluator. Do not interpret this to mean that the evaluator should hold all the power in an evaluative process; rather, the evaluator recognizes who holds power, acknowledges how and why that power is held, and engages with terminology so that power is more evenly distributed. After all, language is a major form of power. For a more specific discussion of language and power in an evaluation, skip ahead to Chapter 13.

BOX 1.2. The Evaluator's Role in Clarifying Terminology

The evaluator's role is to comfortably engage with terminology and, whenever needed, to clarify it with evaluation users and those involved in the evaluation process. To do so, she needs an in-depth knowledge of evaluation, as well as facilitation, negotiation, and education skills.

I am an evaluator for a small organization. I find that sometimes evaluation words are not concrete, and I need to negotiate their meaning with the donor who provides the funding. What if the donor who provides the funding disagrees with what my internal evaluation guide says?

The uses of words that do not have concrete definitions are often determined by those who have the most power. Commonly, though not always, these individuals or groups are those who provide or control the resources or make other key program decisions. If a donor or government agency that provides the bulk of the intervention's funds (and/or is your boss) insists, "This is how we define *X*," this is likely not a good situation in which to have an argument premised on "My definition is right and yours is wrong, and here is the guidebook that backs this up." Rather, what is key is that there is a shared, explicit understanding of how the term will be defined and applied in the evaluation. I strongly suggest that this be written, not just orally agreed upon, and that all people involved agree that *this* is indeed the agreed understanding. The most worrisome thing about providing advice in a book is that there will always be exceptions to the rule; there may be times when bringing your definition, or a different definition, to the discussion is useful. Ask yourself, "How will having this conversation, clarifying this term, and using that definition contribute to a useful evaluation?" And ask the opposite: "If we use that definition, how could it create challenges, and to whom?"

There may be one factor that makes everything a bit tricky: the personal factor. Even if the funding source or management structure does not change (i.e., the government department or donor remains the same), the discussion of terms may still need to be revisited throughout the life of the program or the evaluation. When a new *person* fills the decision-making or funding role, they may bring new interpretations to evaluation terms and processes (often without explicitly stating this), which will leave the recipients confused and often frustrated. The personal factor influences all evaluation processes.

This last point reminds me of another trade secret I must share with you: The personal factor is one of the biggest undiscussed influences in any (and all) evaluative processes. Recognizing, and not ignoring, the fact that individuals influence an evaluation process is critical, as each person can be a tremendous facilitator, an impediment, or just a source of constant bumps and bruises along the way.

I am an evaluator for multiple interventions in a large nonprofit organization. As such, I have multiple donors who all bring different understandings to the same evaluation terms. How do I straighten this all out?

Working in an organization that is accountable to several different donors, departments, or managers, all of whom use the same evaluation terms or concepts in dissimilar ways (or sometimes with just enough nuanced differences), can cause havoc. I recommend starting by clarifying *who* means *what* for each evaluation concept or term, and how each meaning translates into practical application (even for the words with concrete definitions). Then determine in what way practically applying the different definitions and/or interpretations of the same evaluation word influences how to conduct an evaluative process. Compare how the different interpretations will influence what is done; will they look the same in practice? If not, clarify how the differences influence how the intervention will be evaluated and valued, and use that information to have an informed discussion on how to move forward.

Learn the Meaning—Not Just the Label

Evaluators need to use a term or concept with the same understanding as their peers, colleagues, or clients to ensure a useful evaluation. It is with the utmost hesitation that I use the following example to support a conversation about the potential dangers of using labels without having a clear definition of what the terms mean. (Chapter 7 talks about effectively using results labels and refers to the example that follows.) The example is provided to start *our* conversation on labels; it is not a conversation to be used with anyone else on how to use and apply the terms.

Some common labels in evaluation are *input, activity, output, outcome,* and *impact*. Here is one example of how these terms may be used. Some of us monitor how much money is spent to get something (e.g., $100 for running shoes). In evaluation, we call that input, or what was put in to get something out. What the person is doing (e.g., running) is called the *activity*. The very first thing that happens because of that activity, such as how far or how often the person runs every week, is often called the *output*. Some of us may also then monitor how much weight the person loses because he runs so much, which is often what evaluators call the desired or intended *outcome,* or what happens because of the output (e.g., the distances ran). Then, because the person has lost weight, he has higher self-esteem and better overall health, which some evaluators call a *higher-level outcome* (and some call *impact*). He then decides to start a new career and adopts an overall healthy lifestyle, which brings him greater happiness, and some refer to this result as the *impact* (and some call it *longer-term impact*).

This example illustrates a common, simple, and quick way of using these labels without much concern for the dangers of doing so.

I now implore you to be patient and trust our journey. I ask you not to argue with or question what is written in the previous paragraphs (not just yet, anyway), and to not immediately use these labels mechanically as described. Rather, I invite you first to read the next section about the dangers of using labels in evaluation. If you want to jump ahead to read an in-depth discussion on sorting out the input–output–outcome–impact (and so on) discussion, go to Chapter 7 (though I encourage you to keep reading from here through Chapter 10 to see all the linked discussions). After reading *through* to Chapter 10 (or, if you must, after only reading Chapter 7 on results labels), revisit the example that starts on the previous page. You will then understand why I have asked you not to argue (just yet) or to use the labels without a deeper understanding of the questions they pose. I'll see you on the other side.

Learn the Meaning—Not Just the Label

Labels can be useful when we buy a product. For example, "Here is shampoo for babies, and there is shampoo for dogs." Or "That is dog food, and this is cat food." Or "That is poisonous." I buy cat food for my cat, and dog food for my dog. Labels are meant to be helpful. To be honest, however, my dog seems to prefer the cat food, which she always steals from the cat; as such, this provides my first practical example of how perhaps labels are not always helpful in the way we think they should be. Although my dog's eating cat food is expensive for me (as she steals copious amounts), there is little harm done from her perspective (but from the cat's perspective, he may disagree!). It would be much worse if my dog ate something labeled "poison." In M&E, some labels can be just as wily.

Do not get caught up in label-fueled arguments, such as "That is monitoring!"/"No, that is evaluation!" or "That is an impact!"/"No, it is an outcome!" These discussions will exhaust you. If someone uses an evaluative term (even when the word or concept has a concrete definition), ask them what they mean by it. *Descriptions* under the labels are important. When you are engaging with a client, colleague, or friend about an evaluation process or term, the primary goal is for everyone to have the same understanding insofar as what is needed, intended, and expected. Use a label when, and only when, the meaning is explicitly clear to everyone.

I have now talked about the importance of knowing when a word has a *concrete* definition and when it has a *fuzzy* one; about the need to be wary when using evaluation labels; and about the role of the evaluator in bringing clarity to each context. Now let's apply what we have just learned by discussing three convoluted terms in the field.

EVALUATION, EVALUATIVE THINKING, AND MONITORING: WHAT THESE TERMS (CAN) MEAN

The terms **evaluation, evaluative thinking,** and **monitoring** are often used together, and sometimes interchangeably, in an evaluative process. As the next several pages illustrate, there are many "correct" definitions and understandings for each term. However, for many of these words, there is also a non-negotiable core element (or set of elements) in the definition. These elements can be described in multiple ways by various people, implicitly or explicitly, which adds to the trickiness.

Together, let's explore these three terms, and explore how to discuss and engage with them so that their meanings are clear and support straightforward evaluative processes. Facilitation and training tips are provided that can be used to promote self-reflection or to engage others in a discussion aimed at creating understanding and agreement, or simply at clarifying disagreement (the old "We agree to disagree" sentiment). Each is useful in its own way. Let's start with the word *evaluation*.

Evaluation

What is *evaluation*? The word has different meanings and associations for different people, who are informed by, among other things, their education, cultures, values, and exposure to various evaluation experiences. Add to this that organizations, evaluation theorists, social science disciplines, governments, foundations, institutes, and donors often have their own slightly nuanced definitions, and sometimes what appear to be just completely different understandings. Thus asking more than one person, "What is evaluation?" rarely results in a specific and consistent answer. An even more complicated discussion revolves around who is an evaluator, and who is not, as discussed in Chapter 11 of the book. For now, let's focus on uncomplicating the term *evaluation*.

Thomas Schwandt (2015) offers two definitions that provide one place to start a discussion. Schwandt provides a narrow definition of evaluation as a process that draws on research methods to gather information about how a program or policy is working and whether it is effective. He also provides a broader explanation by stating that evaluation is a form of critical thinking that employs evaluation-specific methodologies to judge something. That "something" is broad. I like and will use Schwandt's basic understanding of "something," and I will add a few more throughout this book, some of which are also used by Mathison (2005). When I talk about evaluating "something," I am referring to evaluating an activity, event, intervention, performance, process, product, policy, practice, project, or program. Wait, I am not done yet. Evaluation can also be used to explore and evaluate strategy, systems change, ecological sustainability, resilience, networks, and principles, to give a few additional nouns (Patton, 2015; Rogers, 2016). From this point onward in this text, all of these common nouns will not be written out; this would make the sentences too long. Therefore, when examples are provided, just one of these words is chosen, most often the word *intervention*. Keep in mind from now on that in any example in any part of the book where the word *intervention* is used, any one of these other common nouns can be substituted.

Two further definitions provided by two influential evaluation theorists, Michael Scriven and Michael Patton, involve different perspectives on evaluation, and therefore the definitions differ slightly. According to Scriven (1991), evaluation is the systematic determination of the merit, worth, or value of something, or the product of that process.

BOX 1.3. Merit, Worth, and Significance

Merit refers to the quality of an intervention—the inherent value. *Worth* refers to how others view the intervention, and determining it involves looking at an intervention in its context. For instance, an evaluator has merit if she brings strong evaluation knowledge and solid experience to an evaluation process. She has worth if those involved in the process have found her knowledge and her services useful (Davidson, 2005; Mathison, 2005; Scriven, 1991). It has *significance* if someone considers what the evaluator did to be special or to have particular meaning.

Michael Patton (2008) offers a broader, more multifarious definition, and one that brings in the idea of evaluation's having multiple purposes, uses, and users. He suggests that program evaluation focuses on three elements: "(1) the systematic collection of information about (2) a potentially broad range of issues on which evaluation might focus (3) for a variety of possible judgments and uses" (p. 39). Patton's definition suggests that different people may have different needs for evaluation. These differing needs are likely to result in the use of different questions, approaches, and criteria to value the intervention.

All of the definitions described in this section are widely acknowledged and offer commonly accepted explanations for the role and purpose of evaluation. However, Schwandt's, Scriven's, and Patton's definitions all stem from different perspectives, leading to some nuanced and some blatant interpretations of the term *evaluation.* None of the definitions provided (Schwandt's, Patton's, or Scriven's) contradicts the others per se, but they give very different understandings of evaluation's role in society, and these differences undergird the discussion about what an evaluator does. Nonetheless, at the core of each definition of evaluation, there are two inherent commonalities.

1. *Evaluation is systematic.* Evaluation draws on systematically gathered evidence. Any evaluation process must have clear, transparent, explicit approaches for collecting and analyzing data. The processes of collecting, analyzing, and interpreting the data must be logically consistent (Patton, 2015; Rossi, Lipsey, & Freeman, 2004; Schwandt, 2015).

2. *Evaluation values or in some way judges something.* The word *value* is part of *evaluation.* (Well, sort of, as we do need to add an *e*.) Nonetheless, an evaluation provides more than just findings; it gives a value judgment. Here is where one of the most complicated parts of evaluation often comes in; after all, a value is based on assumptions or worldviews, which are culturally and socially embedded, and very much otherwise context-dependent (Mertens & Wilson, 2012; Patton, 2015; Schwandt, 2015).

There is one more item that is not explicitly stated in these definitions, though it is a core part of defining and discussing evaluation.

3. *Evaluation is political.* A third core element of evaluation is that evaluation is political (Candel, 2018; Chelimsky, 1987; Greene, 2000; House & Howe, 1999; Patton, 2015; Weiss, 1987). What does that mean? Essentially, it means that evaluation can influence—and is influenced by—who gets what, who does not get what, who benefits, who does not, who is even involved, and who is not. Politics are involved in every part of an evaluative process, from the decision to have an evaluation, to how data are interpreted, to how and with whom the knowledge generated from an evaluation is shared, to how decisions are made (Bowman, & Dodge-Francis, 2018; Patton, 2008). See Chapter 13 for a more in-depth look at politics in evaluation.

*Hold on a second. Before you continue, I have a question. The word **data** is used a lot in the explanation. What's the difference between **data** and **information**?*

Data and *information* are words that are sometimes used interchangeably, yet they are not the same thing. Data are the raw facts, figures, and words that, when analyzed to make some sort of sense, become information. Information is shaped or interpreted. In academic language, the word *data* is always considered to be plural. For some reason, this is only true in academic writing. In newspapers, for example, it is acceptable to say "Data is." The word *information* is always treated as singular.

Here is an example of data versus information. My son, Rhys, is 9 years old at the time I am writing this. Some days when I go into his room, he has unceremoniously dumped hundreds (dare I say thousands?) of little colored blocks on the floor. To me, it looks like one big mess. To him, it looks like lots of potential items (data) with which to create (interpret) something. And he does. He creates cars, castles, houses, monsters, and more. He takes what looks like a big mess to me (his data) and creates meaning (information) from it, which he gives to me (his client). To state this another way, when data (the individual blocks) are formed into something through analysis and interpretation (my son's creative design process), they become information.

*Given how you have defined **data** and **information**, then what is **evidence**?*

Evidence supports a statement. So data become evidence when used to support an argument, a hypothesis, or a belief. Evidence comes from data, which most often come from one of our five senses—we can hear it, touch it, smell it, taste it, or see it. And other people can, too. I say "most often," because there are other ways of knowing that do not use these five senses (see Chapter 4 for further explanation). Another way to think about evidence is how it is used in a murder mystery. A detective identifies the murderer with her evidence; the evidence consists of the knife, the body, and the killer's motive.

Because of our acceptance of using some words interchangeably in our daily lives, the boundaries among terms such as *data, information,* and *evidence* can be blurry. Discuss the meaning of *data, information,* and *evidence* (and other similar words, such as *knowledge*) with all parties you are working with or for, and ensure that everyone has a clear and shared understanding of their meaning.

BOX 1.4. Core Elements of Evaluation

There are three fundamental answers to the question "What is evaluation?" Evaluation provides systematic, transparent, and logically collected evidence; it offers a value judgment; and it is political.

Now that we are clear on what data, information, and evidence are, and have looked at some basic definitions of evaluation, let's engage in an in-depth discussion.

Where Evaluation Definitions Begin to Differ

The core understanding we have just discussed is that evaluation is a systematic, transparent process; it values something; and it is political. Even when we agree that these are the core elements, however, differences emerge. These differences are often rooted in who provides the definition (e.g., a person's training and experiences), how evaluation is used in an organization (e.g., its formal definition and practical applications), and its perceived role in society. Thus discussing what is evaluation goes beyond providing a rote definition of evaluation to a discussion that encompasses political, social, and cultural worldviews, among others. Remember, when engaging with colleagues or clients about evaluation, facilitate a conversation that leads to an explicit understanding of what evaluation means to them, what they expect evaluation to do, and what they expect to happen because an evaluation takes place. Do not just ask for their definitions (and do not just provide one).

BOX 1.5. The Meaning of *Empirical*

The word *empirical* is commonplace in evaluation. But what does it mean, exactly? The *Oxford Living Dictionaries: English* (2018) website defines *empirical* as follows: "Based on, concerned with, or verifiable by observation or experience rather than theory or pure logic." I would supplement that definition with the statement that the means of gaining that knowledge need to be transparent and logically consistent.

To encourage learning and dialogue, activities are suggested throughout the book. The first activity described on the next page addresses the term *evaluation*. All activities draw on adult learning principles, which assert that adults' knowledge and life experience should be recognized and drawn upon in any learning opportunity. Drawing on adults' previous experience with conducting, managing, or observing evaluation through their work, or even in their daily lives, will help to ensure that conversations resonate with them, which in turn helps to ground learning. To conduct the activities, facilitation skills are critical. If you are still building those, consider working alongside a skilled facilitator who can demonstrate these skills, and from whom you can learn (i.e., a knowledgeable other).

BOX 1.6. Knowledgeable Others

As we move through the book, you will realize the many skill sets and knowledge areas that an evaluative process may require. When building knowledge and skills as an evaluator, it is often useful to draw on *knowledgeable others*. Susan Tucker, an education evaluation specialist, describes knowledgeable others as people who bring in knowledge and skills that supplement your own knowledge and skill set. These may be knowledge and skills that you have not yet acquired, or never plan on acquiring due to lack of time, differing interests, or other factors. Evaluators often (though not always) work in teams that include knowledgeable others, or subsequently draw in knowledgeable others as needed. In any evaluation process, including knowledgeable others often creates a richer, more diverse team that can bring added value not only through needed knowledge and skill sets, but also through their lived experience and perspectives.

While reading different definitions and explanations of new terms is helpful, having to apply and engage with various definitions concretizes learning. Therefore, the first teaching activity offers a structure for engaging in a discussion centered on the term *evaluation*. If you are new to the field, try applying the exercise with a few colleagues. When you feel more confident, facilitate the activity in a classroom, with clients or colleagues, or with a group of evaluators.

ACTIVITY 1.1. Learning through Discussion: Evaluation

Purpose: The facilitated discussion encourages people to think about how different evaluation definitions can influence a person's understanding, and therefore implementation and expectations, of evaluation.

Time: Allow approximately 5 minutes for preparation, 10 minutes for discussing the two questions, and 15 minutes for group discussion. Total approximate time: 30 minutes. Longer time frames are likely to encourage more in-depth discussions.

Preparation: First, select two evaluation definitions. Second, on flip chart paper, write the selected definition for each evaluation term, and put them on the wall (an option is to provide each group with both definitions). The groups need to see both definitions at the same time. Patton's and Scriven's definitions can be used, or other definitions as appropriate, such as the government department's, organization's, or donor's. Second, write the two facilitation questions, given in the next column, each on their own piece of flip chart paper, ensuring that only the first question can be seen. (Or write both questions on the same piece of paper, but cover the second question.) These two questions will be used during the group discussion.

Preparation of room: Divide the participants into small groups of three to five people, to allow for good discussions. Consider how the groups are organized. You may want to have random groups (e.g., all people with birthdays in January through March in one group, or all people wearing green), or more strategically thought-out groups (e.g., one manager, one evaluator, and one junior researcher). Consider what you want to happen in the groups. Do you want a discussion that mixes various perspectives in one group? Or a discussion among people who are likely to have similar understandings? Do you want one person to learn from another? Or do you want people to build relationships?

The exercise: Once groups are in their own spaces, ask people to silently read the two evaluation definitions. Then ask them to discuss the first question within their groups.

1. In what way are the two evaluation definitions the same? In what ways are they different?

Allow approximately 5 minutes for discussion, and then introduce the second question. It is OK if the group is still talking about the first one.

2. How might these differences influence an evaluative process or an evaluation, with regard to implementation and/or expectations?

Allow approximately 5 minutes for discussion. Then ask the members of one group to share their answers to the two questions, followed by the other groups. Facilitate a discussion that compares the responses. It is important to note that for some groups, facilitating the discussion immediately after the first question, before moving on to the second, may be appropriate; for other groups, that pace may be too slow. Some suggested probing questions during the larger-group discussion include the following:

Which definition do you prefer, and why?

What are the differences and the similarities?

Why is (or what makes) it important to have discussions that lead to shared definitions of evaluation, or at least an understanding that differences of opinion exist?

What might happen in an evaluative process if there was not a shared definition, or agreement to use the same definition, of the term?

Critical learning point: Different evaluation definitions will lend themselves to different expectations, which will likely create challenges during the evaluation. Discussing and developing a clear, shared understanding of evaluation and how the term will be applied are likely to result in shared expectations of the evaluation, which is an excellent place to start any evaluative journey.

Evaluative Thinking

Sometimes people use *evaluative thinking* and *evaluation* in the same breath. While they are indeed closely related, the terms connote different meanings. Evaluative thinking is what encourages the reflection on, and the use of, data and information. It encourages continual thinking about what can be done differently, improved, changed, or enhanced. I have heard some argue that anyone who can think evaluatively is an evaluator. For me, this argument is akin to contending that anyone who knows that it is important to brush their teeth twice a day is a dentist, or that anyone who knows that a steady diet of processed sugar is not good for the body is a nutritionist. Evaluative thinking is a critical part of being a thoughtful and engaged development worker or program manager. Thinking evaluatively means using data and information to be reflective, to learn, and to inform actions. While it is a critical element of being an evaluator as well, evaluative thinking does not by itself make someone an evaluator. (For a rich discussion of why this matters, skip ahead to Chapters 2 and 11.)

Monitoring

Now that we have discussed evaluation, and we have chatted about evaluative thinking, let's tackle the next term often associated with evaluation and evaluative thinking: *monitoring*. Monitoring is a common term used in public management that has crept slowly into evaluation books, journals, and evaluation conferences in the Western world. In the developing or emerging world, monitoring is often a familiar term.

We can all breathe a sigh of relief on how to define monitoring, as different groups tend to agree (though not always) on a common core definition. Monitoring is done on a day-to-day basis, and is the ongoing routine tracking of a program, project, intervention, or policy (or any of the other nouns mentioned earlier in this chapter). The confusion comes in when we are trying to distinguish monitoring from evaluation. Making this distinction is often a thorny conversation in the West, where monitoring can be subsumed into an evaluation process or completely disregarded as not being a part of the evaluation conversation. In developing or emerging markets, monitoring is so closely associated with evaluation that it is difficult to have a conversation that does not reference the M (monitoring) without the E (evaluation) in M&E. Speaking about monitoring as something different and separate from evaluation is practical, however, in that it allows a clearer understanding of what is done when, for what reason, and by whom. Let's take a closer look at monitoring.

Monitoring is something many people do naturally and answers a very basic question: **What is happening?** The question subsumes or includes three foci:

- What is (or is not) happening with the resources used for the intervention?
- What is (or is not) happening in terms of actions to implement the intervention?
- What is (or is not) happening as a result of those actions?

Sometimes we put in a number to gauge if something is happening: For example, did a person run 5 miles or lose 2 pounds? Setting these numbers, which evaluators have often called *targets,* helps us to specifically monitor progress, and show whether the person is moving toward or away from what he set out to accomplish. We discuss setting targets and other related concepts, such as benchmarks, more fully in Chapter 9.

Just a side note here: A critical element to remember in the evaluation journey is that staying on task is critical to untangling and focusing the process. However, because there are so many evaluation concepts and discussions that have interrelated and spin-off conversations, it is easy to get sidetracked. Sometimes a tangential term arises and causes havoc, and there is a need to stop and provide a detailed explanation. At other times, the tangential term adds a needed detail to the conversation (such as *target* above, or earlier in the chapter, when I stopped and explained the terms *data* and *information*) and requires no more than a quick, yet informative, chat. The critical part is, once the tangential term is explained, to return effectively to the core conversation. So let's go back to monitoring.

Let's assume we are monitoring an intervention. Our monitoring data suggest that the intervention is on track and moving toward where it intended to go, or maybe that it is not on track and thus requires one or more management decisions. The management decision could be to make an immediate change, do nothing and continue to monitor, or conduct an evaluation to find out why the intervention is (or is not) on track. The same monitoring data may also be used to demonstrate accountability. For instance, the data may be used in a monthly report that is presented to a board of directors.

Monitoring is a management function mainly used to determine whether an intervention is using appropriate resources and using them appropriately, doing what it aimed to do, and achieving the results it intended. Monitoring can also be used to keep an eye on external or internal factors that may influence the intervention (e.g., the economy, policy shifts, or people). Monitoring is often associated with informed decisions and accountability. Combined with evaluative thinking, monitoring provides a way to use data to learn and reflect; to improve a project, program, intervention, policy, or the like; to identify what was (or was not) achieved; and sometimes to recognize unexpected results.

BOX 1.7. Monitoring

When discussing the definition of monitoring, we need to consider four core concepts: The data collection is (1) empirical and (2) continuous, and data are used (3) to manage the intervention and (4) to provide accountability. Monitoring answers these questions: What is happening, and did the intervention do what we intended it to do? Monitoring does not ask if we are doing the right thing, or if we are doing things right; it also does not probe in any way to find out *why* the intervention did what it said it would do, or *why* it did not. Monitoring does not ask why or place a value on the results. When it starts to venture in that direction, it is more properly called evaluation.

Discussing monitoring is often more straightforward than discussing evaluation. However, monitoring also has its own nuanced definitions and various labels in different countries and in different organizations. For example, some people use the term *performance monitoring,* and others call the same process *performance management*. These terms have specific definitions that refer to tracking performance against stated targets, or goals. Earlier the term *M&E* was introduced, which can involve another fuzzy discussion, so let's talk about that next.

The Siamese Twins: M&E

My daughter, Gemma, has a book about a little girl who is told to count sheep to make her go to sleep (McQuinn, 2010). The little girl lies down and starts to count the sheep. Except the sheep do not behave and stay in a line: They play games, they swim, they hide, and they go in-line skating, all of which makes them very difficult to count. The little girl keeps running after the sheep to try to count them. Finally she exclaims, "You're exhausting!" and, ironically, falls asleep. I feel this way about how *M&E* is tossed around; it exhausts me. When someone says, "I am going to do M&E," I always wonder, "Well, which is it? Is it the M or the E?" Add to this that when I meet M&E experts, and they explain what they do, the explanations often vary considerably from one expert to the next. Oh, and then how the M&E persons define *monitoring* and how they define *evaluation* also often vary (refer to our earlier discussions on these terms). These are some of the conversations that exhaust me, and sometimes help put *me* to sleep.

When separated and defined, each term becomes useful, and then, when they are loosely linked back together, they serve a valuable role. However, when the term M&E is flung around without any questioning of what is being referred to, monitoring and evaluation become the troublesome twins indeed. Let's explore the practical link between the two terms.

The Practical Link

The link between the two terms is very practical. Monitoring an intervention should provide enough relevant data for a program manager (or whoever has the responsibility for the program, regardless of job title) to manage an intervention adequately and be accountable to those with power over the program (e.g., government, donors, community leaders). When something in that monitoring data raises a "why" question (e.g., "Why are people not coming to a training?" or "Why is that neighborhood having better results than other neighborhoods?"), evaluative thinking moves the manager toward evaluation territory, thus linking what is done through monitoring to evaluation. When an evaluator draws on monitoring data during an evaluation, it is important to verify that the data are available, accurate, precise to the extent necessary, and credible. For an in-depth discussion on data and data credibility, please skip to Chapter 5.

BOX 1.8. Monitoring, Evaluative Thinking, and Evaluation

Monitoring data raises the flag; evaluative thinking makes us go "hmmm"; and evaluation answers why and brings the valuing.

Evaluation Triplets: Monitoring, Evaluation, and Evaluative Thinking

While the twins are troublesome, the triplets bring their own mischief. Here is a way to engage with the threesome. Monitoring is a process that answers the question about what is happening (descriptive). When we see that something is or is not happening, and we ask why, that moves us toward evaluative thinking. When we engage in further data collection or analysis processes to answer the why, and bring in a value judgment, that is evaluation. And often the processes associated with these words are all so tangled up into one exacting process that it may be difficult, if not impossible, to separate them easily in practice.

Now that we are near the end of our discussion on monitoring, evaluative thinking, and evaluation, I would be remiss not to introduce a few more acronyms that are becoming more common in the evaluation field. The newer acronyms I am seeing include MEL or MEAL (monitoring, evaluation, and learning), MER (monitoring, evaluation, and research), MERL (monitoring, evaluation, reflection, and learning), and PMEL (planning, monitoring, evaluation, and learning). And there are more. When working for or within an organization, we need to define and discuss their acronyms, what meaning each word brings, and who does what and when for whom, thus starting a conversation that leads to clarity and supports a useful evaluative process.

Alphabet Soup: The Use of Acronyms (or TUOA)

The evaluation field, and often the many technical fields it engages with (e.g., health, education, environment, community safety, climate change), seem to have a love affair with acronyms. In South Africa, nearly every day I hear or see the acronym ANC as an abbreviation for the African National Congress, the current ruling political party (often best known in other countries for being the political party of the late Nelson Mandela). When I started implementing health systems evaluations in South Africa, people talked about ANC all the time. I could not understand why the ANC played such a large role at the community health clinic level. Then one day I realized that ANC, in the health system, most often meant *antenatal clinic*—though sometimes it still meant the African National Congress. I was telling this story to a friend, and she said, "Oh, I thought it meant *antenuptial contract*." Hmm. My first piece of advice? Don't be shy; always ask what an acronym means.

My second piece of advice? Do not use acronyms if you can avoid them. First, it makes texts hard to follow, particularly for those not as familiar with the context as the evaluator or program staff. Second, in some cases, it can lessen the impact or take away the very meaning of what you are trying to convey. For instance, in a report about gender-based violence, using the acronym GBV can remove the very urgency or criticalness of the issue being addressed. It is a very different experience for the brain to read *GBV* rather than *gender-based violence*. A key point of writing an evaluation report (or anything, really) is to engage readers, not distance them.

While the term *research* is more thoroughly discussed in Chapter 2, the three words (monitoring, evaluation, and research) offer another set of triplets habitually intermingled in one conversation. Thus two activities are now provided to engage others in a discussion of these triplets and to provide foreshadowing for the more in-depth conversation about research and evaluation, and researchers and evaluators, in Chapter 2.

ACTIVITY 1.2. Monitoring, Evaluation, and Research Discussion

Purpose: The facilitated discussion brings out how the terms *monitoring, evaluation,* and *research* are often intermingled, and encourages people to think about how the three terms relate to each other in practice and how they are different. The activity works best as an introduction to the three words, not a summary.

Time: Allow approximately 5 minutes for preparation, 5 minutes for a discussion on each word in a small-group format, and 10–20 minutes for a large-group discussion. Total approximate time: 20–30 minutes.

Preparation: With a marker, write the word *monitoring* on a flip chart. On the next (and hidden) sheet of flip chart paper, write the word *evaluation*. Then write the word *research* (also on its own sheet of paper, and hidden). Thus, at the start of the exercise, the participants can only see the first word (monitoring). You may choose to write all three words on the same flip chart paper; if so, however, ensure that only the first word can be seen by the participants.

Preparation of materials: Each group needs three blank pieces of paper large enough to write a two- or three-sentence definition with marker or pen. It is often better to use marker, as then you have the option to attach the participants' work to the wall, and it is readable.

Preparation of room: Organize groups with no fewer than three members and no more than five. This size allows for good discussions. Consider how the people in the room are divided. You may want to have random groups (e.g., all people whose first names start with A–E in one group), or more strategically thought-out groups (e.g., one manager, one evaluator, and one junior researcher per group). See reasoning and hints from Activity 1.1.

The exercise: Before providing any definitions or explanations of monitoring, evaluation, or research, explain to the larger group that it is their job to teach a brand-new colleague (or classmate or staff member) the definition of monitoring. Show the group the word *monitoring* written on a piece of flip chart paper. Provide each group with one piece of paper, and have one person in the group write down the word to be discussed (i.e., monitoring) at the top of this piece of paper. Then ask them to discuss, in their group, how they would explain monitoring to a colleague who is not at this training session. Ask the group to write down the explanation on the paper, below the word. When groups are finished, show the word *evaluation* to them, which is written on the next flip chart. Then have one person in the small group write this word at the top of a second blank piece of paper. Now that they have explained what monitoring is to the new person, ask them to explain what evaluation is to that person. Have groups discuss the term and write down how they would explain evaluation to the same colleague. When this is done, show the word *research* and repeat the steps. After each group has written a definition for each term, have the groups share how they would explain these three terms to their new colleague.

As a facilitator, listen for the distinctions made among the terms (or lack thereof). Ask the groups to provide feedback on their internal discussions, not just their definitions. After this exercise, present the more common definitions (those described earlier in this chapter, or other definitions relevant to the group) to close the session.

When facilitating, consider using these three questions:

- What were the challenges, if any, in describing these words?
- What is the importance of clarifying a distinction among these words? Why does it matter in practice?
- What are some of the likely challenges if the words are not clarified—for example, in common usage, in a job description, at the office, or in an evaluation process?

Facilitation hint: Before the groups present their descriptions of the three terms, tell them that they may only read what is on the cards. Reading what is on the cards helps the groups to focus, and helps keep the activity to its set time.

Optional exercise: To make this activity a bit more fun, use role playing. After all groups have their explanations of the three terms, invite one person to the front of the room to "play" the new colleague. Have the members of each group present their explanations to the person, and then ask the person which one was clearer. There is no intent for a winner; rather, the reasons for choosing one description over the others should invite more discussion.

Critical learning points: Knowing the definitions for these three terms, and practically applying them, are two different things. The group should be comfortable discussing the three different words and how they are practically applied in their organization or in their work (not just providing rote definitions).

ACTIVITY 1.3. Stories about Research, Monitoring, and Evaluation

Purpose: The facilitated discussion brings out how monitoring, evaluation, and research are often intermingled, and encourages people to think about how the three terms relate to each other in practice.

Preparation: None, unless you want to create your own story.

Time: Allow approximately 2 minutes to read the story, 5 minutes for individual thinking time, and 15 minutes for group discussion. Total approximate time: 22 minutes.

The exercise: Read the example. Here is an example that can be used: "When baking a cake, I first **research** what kind of cake I want to make—say, a chocolate layer cake or plain vanilla. I research different recipes, and I choose one. I then set about making the cake, and I **monitor** what I put into it, such as how much flour or salt. I also **monitor** the cake to tell me how brown the top becomes, so that I know when to remove it from the oven. When **evaluating** my cake, I decide if the cake is good or bad, based on my preference, and sometimes I even listen to what others think of my cake." In my example, I have talked about using research, monitoring, and evaluation.

Now ask individuals to develop short examples (one per person) from their everyday life, and provide 5 minutes for the short examples to be written down. Ask for a few participants to share. These are some facilitation questions you can ask:

- What are the practical differences among the three terms?
- What challenges did you have in describing the three different processes, if any?
- How do these critical differences translate to your work?

Critical learning points: Again, knowing the definitions for these three terms, and practically applying them, are two different experiences. The group should be comfortable discussing the three different words and how they are practically applied in their organization or in their work (not just providing rote definitions).

Here are the more common conversations I have with regard to M&E. Reading these dialogues will build your confidence about engaging in these and similar discussions.

If someone holds the job of an M&E officer, or they say they are an M&E expert, what does it mean?

It can mean many things, so take the time to inquire what someone means when they use this term. While some organizations or governments may specify requirements associated with these job titles, there is no accepted or general job description for an M&E officer or an M&E expert. Someone holding this title or position (or self-labeling) may have extraordinary experience and knowledge of M&E, or may have little to no knowledge of either monitoring or evaluation.

Consider when someone takes a position as an M&E expert, and brings substantial knowledge and skills with monitoring to this position, but little to no evaluation knowledge or experience. This creates awkward situations for the M&E expert, who is then asked to do or manage work for which they have no background, training, or experience. At best, the M&E expert recognizes this challenge and seeks advice, training, knowledgeable others, or mentoring. At worst, the person's lack of knowledge and experience damages an intervention, gives a poor name to evaluators and evaluation, and frustrates key stakeholders and others who are negatively affected by a poorly implemented evaluation. So when people say they "do M&E," always ask them, "So tell me, what do you do, exactly?"

Finally, be aware that the same conversation may also happen when someone says they are an evaluator. Being an evaluator can mean many things—as described in Chapters 2 and 11 of this book.

I am not sure if I can ask an M&E person to do an evaluation. What do I do?

Clarify what skills, knowledge, and experience the person has with evaluation. Here are a few questions you can ask to determine the person's likely ability to do an evaluation that will meet your needs:

- What would be the steps that you would use to conduct the evaluation?
- What are your favorite evaluation approaches, theories, or models? Can you tell me about some you have used in the past?
- What are the ways in which you will determine the valuing criteria (i.e., determine their merit or worth), or what processes will be used to value the findings?
- Please tell me about the last evaluation that you conducted (or the most interesting, or the most challenging). What were some of the challenges? What worked well?

Answers to these questions will encourage a dialogue that provides insight into the M&E person's level of evaluation knowledge and experience.

Is a program manager expected to be an evaluator?

While the response to this question will depend on the organization, in general the answer is no. Program managers need "enough" evaluation knowledge to be able to engage with an evaluator, an evaluation unit (internal), or an evaluation team (external or internal), and/or to manage an evaluation. Here is another way to think about answering this question. I do not expect a program manager to be an auditor, or therefore to conduct a financial audit. However, I do expect a program manager to understand his budget, and to know if the intervention is over- or underspending. Similarly, I do expect a program manager to be able to use monitoring data. He needs to know enough to interpret the data and understand what the data are saying—for example, to take action (or not). For instance, in a high school tutoring program, the program manager should be able to answer questions such as these: Do children show up for tutoring (and, if so, how many)? Do children who show up for tutoring have better grades (if that is the intended result)? How many tutors are there who can cover advanced math, and does that number meet the current students' needs? Thus, just as a program manager is not (often) an evaluator, an evaluator is (in most circumstances) not an auditor or a program manager. A program manager needs to understand how to engage with monitoring and evaluation so that it is useful. Please see Chapters 2 and 11 for an in-depth look at the role of an evaluator.

Intervention, Activity, Project, Program: Doppelgangers or Differences?

I like the word *doppelganger*. It has a specific meaning that is exactly what I am trying to convey. Vocabulary.com (n.d.) provides a nice explanation: "Someone who looks spookily like you, but is not a twin, is a doppelganger. Originally, this was a type of ghost. The word doppelganger is German and literally means double walker—as in a ghost or shadow of yourself."

Doppelganger is a wonderful word to describe how the labels *intervention, activity, project,* and *program* are used. For example, *intervention* and *program* are sometimes used interchangeably to discuss the larger picture (e.g., "We have an environmental program," or "We have an environmental intervention"), while *project* is sometimes but not always used to describe a collection of smaller activities (e.g., "Our program focuses on projects that address fish and wildlife conservation"), and *activities* are often specific (e.g., "Our project has several activities, one of which is to issue licenses to fish farmers"). And sometimes not. When you are working with any organization, ask them to describe how they use these labels and terms, so that everyone is clear about what is what, and which label is attached to what description.

Here is an example of how these words may be used in Organization A.

Intervention	Activity	Project	Program
The nonprofit is teaching high school students how to study for exams.	One tutor works with one student for 2 hours per week.	Each site (and there are 450 sites) has five tutors who work 5 days a week for a total of 20 hours.	It is a national program for inner-city youth.

Here is an example of how these words may be used in Organization B.

Project	Intervention	Activity	Program intervention
The nonprofit is teaching high school students how to study for exams.	One tutor works with one student for 2 hours per week.	Each site (and there are 450 sites) has five tutors who work 5 days a week for a total of 20 hours.	It is a national program intervention for inner-city youth.

Vision, Mission, Goal, Objective: Management Speak Translated into M&E Speak

For *mission, vision, goal,* and *objective, doppelgangers* is an appropriate term to reuse here, and let me just say "ditto" to what I have written in the previous section. The management language and the M&E language can be thought of as distant cousins (perhaps second cousins), or we can view them as terms that live in their own separate yet parallel universes. When you are working with management terms, identify what should happen or be achieved when, and to what extent, the organization is held directly accountable for which achievement (i.e., which result). Knowing both (what is intended to be achieved and what the organization is held directly accountable for) will clarify what *needs* to be monitored and evaluated (there are many things that can be monitored, and we delve into this in Chapters 3 and 4), and what is the far-reaching but necessary dream (e.g., a better world, equality for all), which is not (often) assessed in the lifetime of the intervention. I say “necessary,” because most interventions are aiming for something just beyond their reach, and this dream guides the work in the present, like a guiding star. When M&E words are mixed and matched with management words, they tend to (though they do not always) have corresponding levels. It is with trepidation that I note the following common equivalents:

- Vision, mission, goal = Impact
- Objective = Outcomes

Again, these are often somewhat similar levels—but not always. Please jump ahead to Chapter 7 for a more in-depth look at why these examples are provided with some trepidation.

Discussing Concepts, Terms, and Labels

Do not assume that everyone engaging in an evaluation discussion has the same understandings of concepts, terms, and labels—even those that have concrete definitions. Here are three pieces of advice:

- Discuss each concept or term.
- Define and describe how it will be applied.
- Write it down and get physical sign-off on it.

WRAPPING UP

Common evaluation concepts and terms, even when they have concrete definitions, can be defined, interpreted, and applied by different people, groups, and organizations with nuanced or astronomical differences. Knowing the difference between concepts and terms that have specific definitions and uses, and those that do not, is important. An evaluator needs to know when she can say, “That is indeed the definition; let’s discuss how we will use it,” and when she needs to negotiate a concept’s or term’s definition, meaning, and application. When an evaluator ensures that those who need to be on the same page with regard to a term or concept’s meaning are actually in agreement, she avoids mayhem, misunderstandings, frustration, and likely major disappointment in the evaluation and its process. An evaluator who explicitly engages with evaluation words, concepts, and their meanings (whether through a 2-minute conversation, a written clarification, or a longer process) is more likely to have a useful evaluation. Some side benefits of clarifying language (evaluators call these "unintended results") may include building capacity, strengthening relationships among those engaged in the process, and leveling the playing field among stakeholders.

The next chapter moves us from talking about monitoring, evaluative thinking, and evaluation to another topic that is often as intensely discussed: the differences between evaluation and research, and between an evaluator and a researcher. Before you leave this chapter, however, join me in a conversation.

Our Conversation: Between You and Me

Here are two types of situations that you may encounter with regard to the concepts and terms just covered. Practicing with these common scenarios can be a great help as you deal with these kinds of situations in the real world.

1. Your clients (or boss or colleagues) have asked you to do the M&E for a program. You have no clear idea of what they want you to do. Start at the beginning, as we have done in this chapter. Encourage a dialogue to get clarity about what they expect from M&E. To prepare, identify the organization's or client's definition of *M&E* (or its component terms, *monitoring* and *evaluation*), or any description of how the acronym is used in the organization. Reading the organization's information will help you identify a basic place to start asking questions. Here are some general discussion guidelines:

- Be clear about what is within your boundaries and what is outside them, with regard to what you are capable of doing in terms of your own knowledge/skills and the resources that are provided.
- Move the initial conversation away from labels. For example, if the request is "We expect you to do monitoring," a good follow-up question would be "Could you explain a bit more what you mean by *monitoring,* perhaps with some examples?" If they seem hesitant or unsure, offer some suggestions or examples of types of activities often associated with monitoring.
- Seek to understand how the data that you collect, and the processes that you engage in, will be used by whom to do what.
- Once you think you have a clear understanding of what is wanted, provide some concrete descriptive examples of what you are going to do, and have them react to those examples. Provide some examples as well of what you will not do, and have them then react to those examples.

Draw on the definitions, discussions, and exercises demonstrated in this chapter to engage your clients (or boss or colleagues).

2. You are working in a place where English is not the first language, and there is confusion about how to translate or how to use the terms *monitoring, evaluative thinking,* and *evaluation.* A discussion of the distinctions among *monitoring, evaluative thinking,* and *evaluation* is not always a useful conversation when terms are translated. Or the terms may simply not be useful in the culture in which you work, even if English is the dominant language. So remove the labels. Facilitate a process that draws from how monitoring and evaluation are described in this chapter, without using the terms. For instance, ask the person, "How will you know if the intervention is doing what it is supposed to do?" What *is* needed is a clear understanding of what data are collected when, by whom, how, and for what reason, and how those findings will be valued and by whom. Once all that is clear, you can then discuss who does what in that process. Then and only then, choose appropriate labels (e.g., "We will call this *X* and that *Y*"), *if* labels are needed. For example, labels may be needed for the group to engage with their donor or others outside the organization. Or perhaps no labels are needed.

The other day my neighbor said to me, "Tell me again what you do for a living. You are a researcher, right?" Sigh. Please come join me in Chapter 2.