

CHAPTER 1

ISSUES IN ENHANCING EVALUATION USE

John Mayne

Underutilization of evaluation in organizations is much lamented in the evaluation literature (e.g., Mayne, 2009; Pollitt, 2006; Stame, 2004). The literature is full of advice on how to address this shortcoming, which ranges from producing better evaluations (e.g., Patton, 2012; Preskill & Jones, 2009; Russ-Eft & Preskill, 2009) to getting closer to decision making (e.g., Vanlandingham, 2011) to building up an evaluation culture (e.g., Botcheva, Shih, & Huffman, 2009; Hernandez & Visser, 2001; Mayne, 2009; Perrin, 2006) to communicating better (e.g., Ramalingam, 2011). Johnson et al. (2009) provide an overview of the literature on evaluation use. And much of this advice is well founded and useful.

There is also a large body of literature on organizational learning with significant discussion and debate on how organizations learn (e.g., Easterby-Smith, Crossman, & Nicolini, 2000; Garvin, Edmondson, & Gino, 2008). There is also a literature on how evaluation relates to organizational learning (e.g., Torres & Preskill, 2001; Cousins, 2007; Russ-Eft & Preskill, 2009). Cousins, Goh, Clark, and Lee (2004) review much of this literature and make suggestions for enhancing evaluation utilization in organizations.

But a great deal of this and related literature has been written by outside observers, experts, and academics and, as a result, may have to some extent a built-in bias—a bias in favor of focusing on “bad news,” on the problems and shortcomings faced by evaluation in an organization which can then be discussed and analyzed. This perspective is perhaps seen as more interesting reading. Articles on successful evaluation in organizations are much harder to find. Further, those who might be inclined to write these articles such as consultants working for organizations or evaluators inside an organization may see little incentive for writing up “good news” stories—which might have a harder time getting published.

This book’s editors try to counter this phenomenon by getting “insiders” to write up their experiences with evaluation in organizations. The aim is not to have good news stories written, per se, but to explore the actual challenges faced in organizations that are

trying to make evaluation useful and used and to discuss, in light of the related literature, what strategies have been used in practice to address these challenges and enhance the utilization of evaluation.

These insiders are those involved in commissioning, managing, and/or carrying out evaluations in organizations. These individuals and units want the evaluations they are involved with to be useful: to be used to inform decision making at different levels in the organization and lead to improved performance of the organization and its activities. Organizations are not interested in supporting activities that are not adding value to the organization.

The chapters in this book address a variety of issues:

- The different ways evaluation is set up—institutionalized—in government sectors and in public organizations and with what results
- Why it is so hard to make evaluation a regular aspect of good management
- Building organizational cultures that support effective evaluation
- Strategies that are being used to ensure better value for money and enhance utilization of evaluation findings in organizations
- How evaluations can be strengthened to make them more effective

While many of these issues have been addressed in the literature, what this book offers is discussion based on well-grounded experience on implementing evaluations in organizations. The chapters in this book cover a range of institutionalized evaluation, from evaluation being done in and by government agencies (in Canada, New Zealand, Scotland, and Switzerland) to evaluations being done in and by a supranational government (the European Commission), to evaluations being done in and by international organizations (World Health Organization [WHO] and the International Labour Organization [ILO]).

In this introductory chapter, the stage is set for these discussions by setting out some of the underlying issues:

- What is meant by “use” of evaluation?
- Who uses evaluation?
- What are the challenges to making evaluation useful?
- What are the key issues in designing and carrying out evaluations?
- Why are evaluations done, for what purpose?
- How is evaluation structured and organized in an organization?
- What is the context within which evaluation operates in an organization?

The Organisation for Economic Co-operation and Development, Development Assistance Committee (OECD-DAC, 2002) defines (*ex post*) evaluation as “the

systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results” (p. 21). This book adopts a similar broad perspective in that what is being evaluated—the evaluand—can be any grouping of activities aimed at achieving some end in an organization. Thus, evaluation often involves an assessment of some aspect of an organization’s or several organizations’ performance.

Evaluation can also be *ex ante*, an assessment undertaken of a planned project, program, or policy to determine if it is reasonable to expect that the intervention will achieve its intended aims.

DIFFERENT TYPES OF UTILIZATION

Using evaluation can have a range of meanings, all well discussed in the literature (Kirkhart, 2000; Mark & Henry, 2004; Moynihan, 2009; Patton, 2012; Preskill & Torres, 1999, 2000; Weiss, 1998). The following types of use can be identified:

- *Instrumental use*—use made of evaluation findings to directly improve the programming (Caracelli & Preskill, 2000; Nutley, Walter, & Davies, 2007)
- *Enlightenment use*—both long and short term
 - *Conceptual use*—use made of evaluation findings to enhance knowledge about the type of intervention under study, “to influence thinking about issues in a general way” (Rossi, Lipsey, & Freeman, 2004, p. 411)
 - *Reflective use*—use made of the evaluation findings about the program, its operations, and about future strategies, for example, Jacobson, Carter, Hockings, and Kelman’s (2011) discussion on use of workshops to encourage and support reflection on the results of a major conservation evaluation in Australia
- *Persuasive use*—use made to legitimize or criticize an intervention (Rossi et al., 2004, p. 411) or to build support to scale up or out a pilot intervention
- *Process use*—use made of engagement in the process of carrying out an evaluation to better understand both the program and the evaluation process (Cousins et al., 2004; Forss, Rebien, & Carlson, 2002; Patton, 2007, 2008)
- *Symbolic use*—use made of an evaluation to fulfill a requirement to do evaluation or to show support for an intervention area; token use (Patton, 2008, p. 112)

Instrumental use is the kind of use that most people probably imagine for evaluation, especially in organizations that support evaluation. But studies

often suggest that most use is in fact various forms of enlightenment use (e.g., Peck & Gorzalski, 2009). Is this what organizations expect when they fund evaluation activity? Or are they expecting a shorter term return on their investment?

Chapters in this book explore these issues.

USE BY WHOM?

When talking about use, besides being clear about the kind of use or uses that we are discussing, as advisers or evaluators of public sector performance, we also need to be clear about just whom we are referring to. There are many potential users of evaluative information, both inside and outside an organization, such as the following:

- Insiders
 - Program managers
 - Policy analysts
 - Senior managers
- Outsiders
 - Politicians
 - Control bodies
 - Stakeholders
 - Academics
 - Critics
 - Beneficiaries

These different users will have different

- perspectives on what is being evaluated in terms of knowledge about and experience with the program and its history, views on its relevance and importance, and so on;
- perspectives on evaluation, from seeing evaluation as a tool for enhanced understanding of the program and how it is operating to a tool to provide critical evidence on the program;
- vested interests in the program and its future from unquestionable support to outright hostility or skepticism;
- information needs about the *what* that is being evaluated that the evaluation could address (Mayne, Divorski, & Lemaire, 1999, pp. 28–29).

This simple delineation of different types of uses and different potential users of evaluation begins to illuminate the complex nature and meaning of the seemingly simple issue of “enhanced use of evaluation.”

WHY EVALUATIONS MAY NOT BE USED

Authors discuss many reasons in the literature about why evaluations may not get used in organizations (Nutley et al., 2007; Torres & Preskill, 2001):

Poor product quality (Schwartz & Mayne, 2005)

- *Not timely*—evaluations that do not produce findings when they are needed for decision making
- *Not relevant*—evaluation findings not related to the decisions being made
- *Not credible*—evaluation findings, conclusions, and recommendations not seen as credible enough to be used in decision making (weak data and/or analysis, biased)
- *Not clear*—evaluation findings not well communicated, with no communication strategy and no recognition of different stakeholders’ interests

Poor evaluation process (Johnson et al., 2009)

- *Not inclusive*—evaluation considered the purview of the “experts” only

Limited evaluation culture

- *Not accepted*—evaluation findings not given much weight in comparison to experience and wisdom, unless perhaps they support the experience

Unclear evaluator role

- Evaluator seen more as an outsider, as an auditor, and hence seen as a burden, not seen as useful

Uncomfortable findings

- Program managers, senior management, and/or politicians who may not like or agree with the findings of the evaluation, which may challenge their beliefs or agendas

Most of these “challenges” can be a death knell to utilization, or at least some types of utilization. That is, any one could well result in limited use. It is evident

that a lot can go wrong. Many diverse factors have to come together to make evaluation work well, that is, get used in organizations. This might explain why it is so difficult to have good utilization of evaluation in an organization; many factors have to line up. The chapters in this book discuss these kinds of issues.

EVALUATION FOR WHAT PURPOSE?

Another important aspect to consider is the intended purpose of the evaluation. A useful distinction can be made between evaluation done for learning and that done for accountability. Immediately, one can see the implication for use and by whom.

Much evaluation seems to be done or is seen as being done for accountability purposes (Mayne, 2007). That is, it is done for someone outside the organization (or outside the evaluand) to check on how well things are going, typically to assess if the program should continue to be funded. Most international aid donors, for example, require evaluations as a condition of continued funding. As a result, use by the evaluand, or organization, for improving design or delivery is much less likely, with attention focused more on defending the program than learning what is working and what is not. Evaluation with a clear goal of learning about programs and how they are working will have a much greater chance of being used by the evaluand.

But need there be such a dichotomy between evaluation for learning and evaluation for accountability? Some have argued (Auditor General of Canada, 2002; Mayne, 2007; Perrin, 2007; Zapico-Goni, 2007) that in the age of accountability for results—which programs do not have complete control over—and especially in the context of increasingly complex interventions, accountability should in large part be focused on accountability for learning rather than for results per se. This could be a way of reducing the accountability versus learning tension.

In any event, how organizations see their evaluation efforts is clearly quite important.

KEY EVALUATION DECISIONS

The quality of an evaluation product and process is largely determined by how the evaluation is designed and carried out. An organization undertaking an evaluation needs to make quite a few important decisions. Mayne (2011) provides a list:

- Which program to evaluate?
- What evaluation process to follow? Whom to involve in the evaluation?

- Which issues about the performance of the program to address?
- How credible does the evidence have to be?
- How much time and resources to spend on the evaluation?
- How to make decisions on the methodology (evaluation design, data gathering, and analysis)?
- Which evaluators to use to carry out the evaluation?
- How evaluation findings, conclusions, and recommendations receive agreement or approval?
- What gets reported to whom and when, during and at the end of the evaluation?
- How to determine the terms of reference for the evaluation?

The last issue on terms of reference—what the evaluation will comprise—is where many of these decisions are manifest. Who decides the terms of reference can decide most of the key issues in an evaluation, although it clearly depends on how comprehensive the terms of reference are.

How and by whom these decisions get made will affect the quality of the evaluation product and process. All are potential sources of bias in an evaluation. How such questions get decided determines how an evaluation study or the evaluation function in an organization is structured and organized.

These decisions could be made by any or all of the following:

- Those calling for—commissioning—the evaluation
- Those funding the evaluation
- Those managing the evaluation
- Those managing the program being evaluated
- Stakeholders associated with the evaluation
- Those carrying out the evaluation—the evaluators
- The policy under which evaluation operates

EVALUATION STRUCTURES IN ORGANIZATION

Wide variation exists on how evaluation is structured in organizations, as is made clear in the various chapters in this volume. This is not unexpected given the wide variation among organizations in how they are set up and managed:

- Significantly sized internal Evaluation Units that conduct much of the evaluation activity of the organization, as well as minimally sized internal organization units, which rely on externally hired evaluators to do most of the evaluation effort

- Evaluations commissioned within the organization and evaluations commissioned from outside the organization by, for example, governing bodies
- Organizations with a formal evaluation policy and structured practices and those with a more informal approach to evaluation
- Organizations with well-established performance-monitoring practices and those with limited monitoring efforts
- Organizations with Evaluation Units organizationally quite separate from programming and organizations with evaluation embedded in programming with no central Evaluation Unit per se
- Organizations with Evaluation Units strongly supported by senior management and organizations with Evaluation Units with a less strategic support

Given such wide variation, it is clear that the challenges faced by evaluation in organizations and responses to those challenges will also vary considerably, providing a wealth of potential discussion, as the chapters in this book demonstrate.

In particular, the second introductory chapter in this book, “Evaluator, Evaluand, Evaluation Commissioner: A Tricky Triangle,” discusses the issues of who decides what in relation to the different possible structures between the evaluators, the program being evaluated, and those commissioning the evaluation.

THE ORGANIZATIONAL CONTEXT

The history, nature, and values of the organization in question also clearly will have an impact on the extent to which it uses evaluation. Some organizations have a long history of evaluation activities, others, a much more limited experience.

More fundamentally, the organizational culture matters greatly, in particular the extent to which they have an evaluative culture. I have described a strong evaluative culture in an organization (Mayne, 2009, p. 4):

- Engages in self-reflection and self-examination (i.e., self-evaluation)
 - Deliberately seeks evidence on what it is achieving, such as through monitoring and evaluation
 - Uses results information to challenge and support what it is doing
 - Values candor, challenge, and genuine dialogue

- Engages in evidence-based learning
 - Makes time to learn
 - Learns from mistakes and weak performance
 - Encourages knowledge transfer
- Encourages experimentation and change
 - Supports deliberate risk taking
 - Seeks out new ways of doing business

Implied in these characteristics of a strong evaluative culture, although not explicit, is a willingness to demonstrate publicly what the organization is accomplishing and a willingness to report on its performance and tell its performance story. Thus, added to this list describing a strong evaluative culture should be the following:

- Encourages public reporting on its performance
 - Makes evaluation reports public
 - Reports regularly on the performance of its activities

Authors of many chapters in this book recognize the importance of building and strengthening the evaluative culture in an organization and discuss their experiences in doing so.

Overall, we see a wide variety of issues and contexts that can influence evaluation utilization in an organization. This perhaps should not be unexpected given the complex nature of organizations and how they work and make decisions. Evaluation is one tool for trying to bring evidence to inform such decision making.

CHAPTERS IN THIS BOOK

The authors of the chapters in this book offer insights gained from hands-on experience in the challenges faced in making use of evaluation in an organization. They in general address the following questions:

- What is the context in the organization?
- What specific challenges were faced?
- What does the literature say?
- What actual strategies were used in addressing the challenges?
- How well did the strategies work?

The chapters are written by people closely involved with evaluation in the relevant organizations who share their experiences and lessons learned in the ongoing effort to make evaluation useful to the organization.

The book brings attention to those who call for and commission evaluation. In the second chapter of the book, in addition to the need to consider those being evaluated and those doing the evaluation, Bastiaan de Laat argues for the need to focus on evaluation commissioners, a focus that has received considerably less attention in the literature. He then proceeds to discuss the relationships among these three parties: the evaluand, the evaluator, and the commissioner in different evaluation situations, de Laat's "tricky triangle."

The next five chapters are discussions on a variety of cases of evaluation being done for a government department, agency, or ministry. Penny Hawkins leads off with a discussion on two key aspects of commissioning evaluations: the need to be able to manage the evaluation process and the need to be able to intelligently take account of context. She discusses a number of challenges and good practices in managing evaluations, based on her long experience in the New Zealand Ministry of Foreign Affairs and Trade and the Ministry of Social Development and her more recent experience with the Rockefeller Foundation. She then discusses in particular the importance of culture and context in evaluation, based on experiences in conducting evaluations with the Māori of New Zealand, emphasizing the need to diligently take into account the indigenous people's own culture and social practices, some of which do not immediately fit with "standard" evaluation practice.

Marlène Läubli's chapter on the ups and downs over several decades of building an Evaluation Unit in the Swiss Federal Office of Public Health illustrates the constant need to build and maintain evaluation in an organization. The need to be able to adapt to changing organizational contexts and realities and the challenges of building an evaluation culture in the agency are highlighted. She asks whether institutionalization of evaluation helps or hinders the development of an evaluation culture. Numerous lessons for evaluation commissioners and evaluators are discussed, such as the need for good communication, deliberate planning at the outset for use of findings, and the need to educate programs and commissioners about evaluation.

In her chapter, Erica Wimbush discusses efforts in the public health field in Scotland to bring evaluation evidence to bear on decisions about public health issues. She discusses first the literature on knowledge utilization and the emergence of "knowledge to action strategies" as ways of bridging the gap between the evaluator-researcher and the policy decision maker. Wimbush introduces the context within which Scotland's public health agency operates and describes several strategies the agency has been working with to enhance the use of

evidence. A number of lessons learned are highlighted, such as the need to create reflective space for the consideration of available evidence, the usefulness of identifying champions, and the need for collaboration and networking.

Nancy Porteous and Steve Montague discuss the evolution of internal evaluation in the Canadian Public Health Agency, shifting from the traditional focus on one-off evaluation studies of individual programs to taking a more holistic, organizational approach to evaluative work in the agency. The chapter indicates four strategies they used for this transformation: keeping well informed about the organization, creating an organizational-level logic model, developing a multiyear evaluation plan for the organization, and focusing on cumulative evaluative learning across the organization. The result, they argue was not only improved quality and use of evaluation but also an enhanced ability of the organization to learn and adapt.

The chapter by de Laat and Williams on evaluation at the European Commission (EC) presents a case of a supra-level government organization and reports on the findings from two large empirical studies undertaken on evaluation at the EC. Five types of use were identified and a large number of evaluations reviewed to see what use had been made of them. Based on those findings, the implications for commissioning evaluations are drawn out. The chapter concludes with a discussion of a number of factors that are seen as affecting use, such as the timing of the evaluation, senior management support, involving potential users, and following up on recommendations.

The next two chapters of the book show evaluation in international organizations, namely, the UN system. The World Health Organization (WHO) is a well recognized UN body and a typical UN organization with both a “centralized” corporate Evaluation Unit and a considerable body of “decentralized” evaluation being conducted around the world in its regional and country offices. After describing this complex evaluation environment, Maria J. Santamaria Hergueta, Alan Schnur, and Deepak Thapa discuss the evolution of evaluation in WHO over the past several decades that involve first being located in planning, then with the oversight services, and now facing financial pressures. Challenges to developing and building an evaluation culture within WHO and an outline of a new evaluation policy to address these issues are presented. Also discussed are the pros and cons of locating the evaluation within the Internal Oversight Services, which undertakes audits and investigations.

The second UN case is evaluation at the International Labour Organization (ILO). Janet Neubecker, Matthew Ripley, and Craig Russon argue the need to measure the extent and nature of the use of evaluation findings and present a five-level utilization maturity model, which they developed at ILO. They discuss how specific strategies can be developed to enhance utilization at each

level at each stage of the programming model and by each of the partners in the evaluation process. To build a culture of inquiry, they discuss the importance of establishing and maintaining an active evaluation network of evaluators and managers. Criteria for quality evaluation reports, recommendations, management responses, and dissemination of reports are outlined. When and how each partner in the evaluation process can enhance utilization is discussed.

The concluding chapter looks across the specific cases discussed and draws out themes and lessons that have been identified. Collectively, a wealth of experience is presented in the chapters and considerable practical advice offered on how to enhance use of evaluation in organizations.

REFERENCES

- Auditor General of Canada. (2002). Modernizing accountability in the public sector. In *December 2002 Report of the auditor general of Canada to the House of Commons* (Chapter 9). Retrieved from http://www.oag-bvg.gc.ca/internet/English/parl_oag_200212_09_e_12403.html
- Botcheva, L., Shih, J., & Huffman, L. C. (2009). Emphasizing cultural competence in evaluation: A process oriented approach. *American Journal of Evaluation, 30*(2), 176–188.
- Caracelli, V. J., & Preskill, H. (Eds.). (2000). *The expanding scope of evaluation use: No. 88. New directions for evaluation*. San Francisco, CA: Jossey-Bass.
- Cousins, B., Goh, S. Clark, S. & Lee, L. (2004). Integrating evaluative inquiry into the organizational culture: A review and synthesis of the knowledge base. *Canadian Journal of Program Evaluation, 19*(2), 99–141.
- Cousins, J. B. (Ed.). (2007). *Process use in theory, research, and practice: No. 116. New directions for evaluation*. San Francisco, CA: Jossey-Bass and American Evaluation Association.
- Easterby-Smith, M., Crossman, M., & Nicolint, D. (2000). Organizational learning: Debates past, present and future. *Journal of Management Studies, 37*(6), 783–796.
- Forss, K., Rebien, C. C., & Carlson, J. (2002). Process use of evaluations. *Evaluation, 8*(1), 29–45.
- Garvin, D. A., Edmondson, A. C., & Gino, F. (2008, March). Is yours a learning organization? *Harvard Business Review, 109–116*.
- Hernandez, G., & Visser, M. (2001). *Creating a culture of inquiry: Changing methods—and minds—on the use of evaluation in nonprofit organizations*. James Irvine Foundation. Retrieved from http://www.irvine.org/assets/pdf/pubs/evaluation/Creating_Culture.pdf
- Jacobson, C., Carter, R. W., Hockings, M., & Kelman, J. (2011). Maximizing conservation evaluation utilization. *Evaluation, 17*(1), 53–71.
- Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation, 30*(3), 377–410.
- Kirkhart, K. E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation, 88, 5–23*.

- Mark, M., & Henry, G. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10(1), 35–57.
- Mayne, J. (2007). Evaluation for accountability: Reality or myth? In M.-L. Bemelmans-Videc, J. Lonsdale, & B. Perrin (Eds.), *Making accountability work: Dilemmas for evaluation and for audit* (pp. 63–84). New Brunswick, NJ: Transaction.
- Mayne, J. (2009). Building an evaluative culture in organizations: The key to effective evaluation and results management. *Canadian Journal of Program Evaluation*, 24(2), 1–30.
- Mayne, J. (2011). Independence in evaluation and the role of culture [Abstract notes from a chapter]. *Evaluation Notes, No. 1: INTEVAL*. Retrieved from <http://www.inteval-group.org/IMG/ckfinder/files/Inteval%20Notes%202011-1.PDF>
- Mayne, J., Divorski, S., & Lemaire, D. (1999). Locating evaluation: Anchoring evaluation in the executive or the legislature, or both or elsewhere? In R. Boyle & D. Lemaire (Eds.), *Building evaluation capacity: Lessons from practice* (pp. 23–52). New Brunswick, NJ: Transaction.
- Moynihan, D. P. (2009). Through a glass, darkly: Understanding the effects of performance regimes. *Public Performance & Management Review*, 32(4), 592–603.
- Nutley, S. M., Walter, I., & Davies, H. T. O. (2007). *Using evidence: How research can inform public services*. Bristol, UK: Policy Press.
- Organisation for Economic Co-operation and Development, Development Assistance Committee (OECD-DAC). (2002). *Glossary of key terms in evaluation and results based management*. Paris, France: Author. Retrieved March 8, 2008, from <http://www.oecd.org/dataoecd/29/21/2754804.pdf>
- Patton, M. Q. (2007, Winter). Process use as a usefulism. *New Directions for Evaluation*, 2007(116), 99–112.
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2012). *Essentials of utilization-focused evaluation*. Thousand Oaks, CA: Sage.
- Peck, L. R., & Gorzalski, L. M. (2009). An evaluation use framework and empirical assessment. *Journal of MultiDisciplinary Evaluation*, 6(12), 139–156.
- Perrin, B. (2006). *Moving from outputs to outcomes: Practical advice from governments around the world*. Washington, DC: IBM Centre for The Business of Government and World Bank. Retrieved from <http://www.businessofgovernment.org/report/moving-outputs-outcomes-practical-advice-governments-around-world>
- Perrin, B. (2007). A new view of accountability. In M.-L. Bemelmans-Videc, J. Lonsdale, & B. Perrin (Eds.), *Making accountability work: Dilemmas for evaluation and for audit* (pp. 63–84). New Brunswick, NJ: Transaction.
- Pollitt, C. (2006). Performance information for democracy: The missing link? *Evaluation*, 12(1), 38–55.
- Preskill, H., & Jones, N. (2009). *A practical guide for engaging stakeholders in developing evaluation questions*. Princeton, NJ: Robert Wood Johnson Foundation Evaluation Series. Retrieved from <http://www.rwjf.org>
- Preskill, H., & Torres, R. T. (1999). *Evaluative inquiry for learning in organizations*. Thousand Oaks, CA: Sage.
- Preskill, H., & Torres, R. T. (2000). The learning dimension of evaluation use. *New Directions for Evaluation*, 2000(88), 25–37.
- Ramalingam, B. (2011). *Learning how to learn: Eight lessons for impact evaluations that make a difference*. London, UK: Overseas Development Institute. Retrieved from <http://www.odi.org.uk/resources/download/5716.pdf>
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage.

- Russ-Eft, D., & Preskill, J. (2009). *Evaluation in organizations: A systematic approach to enhancing learning, performance, and change*. New York, NY: Basic Books.
- Schwartz, R., & Mayne, J. (2005). Assuring the quality of evaluation: Theory and practice. *Evaluation and Program Planning*, 28(1), 1–14.
- Stame, N. (2004). Theory-based evaluation and varieties of complexity. *Evaluation*, 10(1), 58–76.
- Torres, R. T., & Preskill, H. (2001). Evaluation and organizational learning: Past, present and future. *American Journal of Evaluation*, 22(3), 387–395.
- Vanlandingham, G. (2011). Escaping the dusty shelf: Legislative evaluation offices' efforts to promote utilization. *American Journal of Evaluation*, 32(1), 85–97.
- Weiss, C. H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), 21–33.
- Zapico-Goni, E. (2007). Matching public management, accountability and evaluation in uncertain contexts. *Evaluation*, 13(4), 421–438.