Reviewed by *Linda A. Pursley*
Lesley University
United States

As part of the series "Evaluation and Society," this edited volume brings together thirteen contributors with extensive experience to share perspectives on international development evaluation. The Rockefeller Foundation provided a grant to Claremont Graduate University to convene this group of leaders in global evaluation, who were then asked to share their papers describing best practices, approaches and principles guiding evaluation work in seven outcome areas. The publication is "intended to enhance learning in development and philanthropy by advancing concepts that can guide practitioners and support understanding in international development evaluation"(p. 2).

The editors, Donaldson, Assam, and Conner, contribute an introductory chapter, "Searching for Approaches to Improve International Development Evaluations," which contains a brief overview of each of the remaining chapters. Their concluding chapter, "Future Directions for Improving International Development Evaluations," includes an appendix of guidance notes from each of the chapter authors, highlighting key issues in each outcome area. Since each chapter has a distinct focus, this
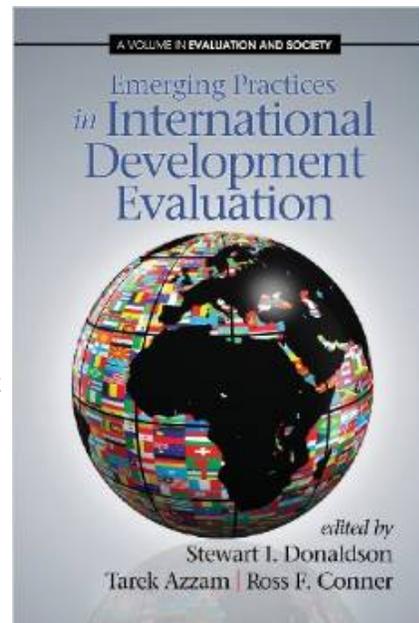
review will address them individually and conclude with observations about the volume as a whole.

Ofir and Kumar set the stage in the second chapter for the subsequent chapters that focus on outcome areas, by articulating some of the features that distinguish evaluation in the context of developing countries, and reasons why they believe that many evaluations are not as useful or responsive to local stakeholders as they could be. Systems and societies in developing countries are characterized as less predictable and stable, often impacted by forces of foreign intervention and globalization that pressure them to adopt external definitions of "development" and "success." Most people lack resources and choices, which can lead to feelings of inadequacy and marginalization. Within these societies, there are generational and power differentials that evaluators must understand to implement evaluations that will be useful and used. Key dimensions on which organizations in developed and developing countries differ are summarized in a table.

The authors discuss three sets of values or frameworks commonly used to focus development evaluations: human development, human rights, and human security, and conclude that in development contexts evaluations must ultimately address the extent to which any initiative contributes to "improving the quality of people's lives, promoting equity, enhancing security, reducing discrimination, and empowering people" (p. 14). In another table, they identify six key issues, assumptions and illustrative questions about evaluation approaches that can be used to "apply a developing country lens." While these questions may be difficult to answer, they are worthy of consideration by any evaluator in international development.

In Chapter Three, Lusthaus and Rojas argue that the evaluation community has focused on policies, projects, and programs to the exclusion of the organization. Further, they contend that since donors are increasing their investment in organizations and institutions, the need to understand their performance is increasing, so they recommend linking evaluation of organizational performance to the evaluation of the programs and policies it implements. They provide a brief overview of the history of organizational performance assessment (OPA), summarizing the evolution of definitions over time. They present three categories of models or frameworks for OPA based on what they seek to identify: best practices or

standards, relationships between concepts or variables, and metrics and results, providing specific examples of each. These chapter authors also provide an overview of the steps involved in conducting OPA, and make recommendations for insuring that OPA results are useful and used. Finally, they conclude with a discussion of the following conceptual and methodological challenges to increasing the use and value of OPA: lack of consensus on the importance of organizations in solving social problems and the need to evaluate them, lack of standards or tools to assess complex organizations, lack of concepts to describe accountability in not-for-profits, need to adjust to new organization types (such as coalitions, networks, public-private partnerships,) increasing concern for organizational social responsibility, and lack of valid data to respond to questions related to organizational performance. The authors mention the need to consider context, but more detail on what this might mean in a developing country would have been instructive, since most of these concepts, models and frameworks appear to have originated in more developed countries. Appendices provide specific examples of OA questions and indicators, as well as suggestions for integrating gender equity into OPA.

In his chapter on evaluating capacity development, Morgan maintains that although it is possible, we are some distance from that goal at present. He identifies a lack of clarity about the definition, which leads to a lack of consensus about what constitutes evidence of capacity development. The evaluation task is further complicated by the fact that some definitions include a "potential state" as part of capacity, and by the complexity and multidimensional nature of the change that is always taking place in organizations seeking to build capacity. For the purposes of his chapter, he defines capacity as "that emergent combination of individual competencies and collective capabilities which enables a human system to create value for others" (p. 77,) and capacity development as "The process of change that, both intentionally and indirectly, contribute to the emergence of capacity over time" (p. 78).

Morgan asserts that conventional logic models are less appropriate for complex adaptive systems and favors "development evaluation" as an approach that combines monitoring and evaluation. He argues against indicators that he characterizes as most useful in determining "what and when" and in favor of measures that give insight into

the "how and why" needed to understand capacity development. It is important that evaluators and international development agencies (IDAs) seeking to build capacity recognize that failure is inevitable, and that in the context of a broader learning process failure may even be a precursor to future success. In his view a key objective of any capacity development evaluation should be building local capacity to carry it out, which means that engagement of local participants is critical. At the same time, he notes that competent local colleagues may be stretched very thin, and cautions that at times, participation can place them in physical or professional danger. Appendices to this chapter include: Definitions of Capacity, Evaluation Methodologies, Some Barriers to Effective Capacity Evaluation, Types of Capacity Development Interventions, and Positioning and Capabilities of the IDA.

In Chapter Five Carden and Duggan focus on Evaluating Policy Influence, the ways in which research impacts policy. Citing Weiss (1977, 1982), they note that often the use of information may be more symbolic and long-term than immediate and instrumental, an indirect influence they describe as expanding policy capacities for both researchers and decision makers. They base their chapter on the Canadian International Development Research Centre (IDRC) analysis of twenty-three cases, which sought to define policy influence, determine what evidence of influence was available, and identify key factors facilitating or inhibiting policy influence.

This analysis revealed that context was critical and five variables played central roles: capacity of policy makers to apply research, stability of decision making institutions, decentralized vs. tightly controlled governance, economic conditions, and state of country transition. Within the decision context, the presence or absence of interest in the problem, demand for information, leadership, and resources on the part of both researchers and policymakers were articulated as elements in a framework that determined the potential for policy influence. The analysis also identified the following steps that researchers can take to influence policy: develop strong communication skills, establish networks of influence, and build relationships and institutions. They recommend a realist approach (Pawson & Tilley, 1997; Pawson, 2006) to the evaluation of policy influence based on the fact that it is one of the few approaches that includes context within the evaluation framework, which they see as central to

understanding policy influence. A figure helps to illustrate the relationships between the research intervention, overall and decision contexts, facilitating mechanisms, and outcomes. Their analysis positions development researchers as change agents, and offers evaluators an opportunity to consider how they can increase the likelihood that sound research will influence policy.

In Chapter Six, Creech focuses on Evaluating Networks and Partnerships. She relies on her experience in the Canadian-based International Institute for Sustainable Development (IISD) in managing knowledge networks for sustainable development. Her chapter describes lessons learned as well as challenges to evaluation practice. Acknowledging the multiplicity of terms used in relation to collaborative efforts, she identifies three forms of collaboration: networked governance, networks of individuals, and inter-organizational relationships (IORs). She observes that the internet and social media are driving the current shift from collaboration among institutions, to individuals connecting to share knowledge and address common concerns, but due to the proliferation of IORs she elects to focus on the challenges in evaluating these relationships. Based on her IOR evaluation practice, she observes that evaluators are often called upon to use the Organization for Economic Cooperation and Development (OECD) criteria: relevance, effectiveness, efficiency, and sustainability, with little clarity on how they might apply and what might serve as performance indicators. In addition, most IORs lack both the culture and tools for self-monitoring and assessment. She maps the OECD criteria to questions that might be asked in the evaluation of an IOR, and specifies "what to watch for" when following these criteria. Her experience suggests that it is important to focus on the change that the IOR seeks to achieve (external value), the operating model for the collaboration, the presence and fostering of social capital, and the presence or absence of critical success factors with in the IOR. Appendices provide an example of guidelines for applying OECD criteria, and list selected IISD consultancies to partnerships and networks.

Evaluating coalitions is the focus of Chapter Seven by Rayno, who situates coalitions within the area networks and partnerships. He places them on two dimensions, formality and accountability, asserting that coalitions are generally more formal and accountable than networks, and less formal and accountable than strategic alliances. He

suggests that coalitions are comprised of six different sets of potential relationships, between: coalition and external world, organizations and the coalition, individuals and the coalition, individuals and the organization, individuals to individuals, and organization to organization, relationships that could be evaluated in terms of their contribution to the effectiveness of the coalition. His chapter offers an example of a logic model for evaluating an advocacy initiative that uses a coalition as a strategy. With respect to evaluation criteria, he suggests and specifies measures in three areas: capacity of organizations to be good coalition members, capacity of the coalition to carry out their activities/strategy, and the short and long term outcomes/impact of the work of the coalition. He suggests that mixed methods are generally most appropriate, and that social network mapping has potential to contribute. The chapter concludes with a focus on the challenges in evaluating coalitions in the context of development, including significant power differentials among partners, difficulty in sharing highly context specific results, sometimes competing demands of funders, and the potential politicization of coalition work.

In Chapter Eight, Evaluating Sustainable Development, Bass and Bradstock assert that inadequate monitoring and evaluation of outcomes are the key reason why increased commitments to sustainable development (SD) have resulted in limited evidence of progress, though more somewhat more progress may be evident in enabling conditions (outcomes) than sustainability impacts. At the same time they believe that strong monitoring and evaluation can document progress, as well as the need to change direction on what they characterize as a "journey" toward more SD. They identify multiple objectives, multiple levels, varying time scales, and the uncertainty inherent in complex systems as challenges to evaluation of progress toward SD. Their brief overview of the history of SD credits the 1987 Bruntland Commission for the most common definition, "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (p. 166), while also characterizing this definition as "ambiguous." Referencing several influential international conferences, they conclude that recent gatherings have confirmed a framework for SD that considers economic development, environmental protection and social progress, while observing that this framework fails to explicitly address governance and

accountability. However, to date the plans, tools and models put forward have not generated the change needed for SD, and they suggest that evaluation could be key. Evaluation could raise questions about local, traditional and non-Western approaches to SD, and help to identify improved incentives, criteria, and accountability mechanisms. These authors enumerate eight challenges in evaluating SD, and then suggest specific measures for assessing both process and impacts. Three common measurement approaches are discussed: indicator-based, reference frameworks (e.g., OECD Principles) and narrative approaches (stories of success/failure.) Learning groups, peer reviews, external auditing, watchdogs, public expenditure reviews, outlook studies, and public engagement are proposed as the best available methods for evaluating SD. They suggest that strategic monitoring is critical to successful evaluation, while observing that the lack of data and/or cost of data collection has been and remains a major barrier in many countries. They propose that another approach to SD evaluation might focus on the extent to which the underlying causes of Unsustainable development or "disabling conditions," have been addressed. The chapter appendix suggests a comprehensive framework for monitoring SD strategies, to provide useful evaluation data.

In the final chapter focused on outcome areas, Rochlin and Radovich write about Evaluating Innovation, specifically innovation for development, citing examples such as micro-finance, use of mobile phones, and community cooperative services. They note that definitions of innovation vary widely and may include products, processes, markets and organizational methods, and may be categorized as incremental, additive, complementary or breakthrough. They define innovation for development as "an invention that is applied and adopted at scale" (p. 201), and further suggest that innovations for development "create economic, political, social, or environmental transformations that seek to improve the health and well-being of poor people and ecosystems" (p. 201-202). Listing characteristics of innovative organizations, they stress the importance of leadership, organizational processes, focus on the end-users, embracing and learning from failure, and the use of diffusion and collaboration to facilitate adoption. According to these authors, failure to include women/gender analysis, over-focus on a specific output or product, and longer timeframes are among the challenges to

innovation for development. The chapter concludes with seven considerations for evaluating grant making in innovation for development, and provides an example of comprehensive model designed for an actual grant maker, which details key performance areas, questions and sub-questions, indicators/evidence, and data sources and methods.

It is possible to identify a number of common themes and connections among the chapters comprising the book. In Chapter Ten the editors identify a search for common definitions and advocacy for mixed methods as common themes. While the variety of definitions for many of these outcome areas may enable the evaluator to select the most appropriate for the evaluation focus and context, this range of different definitions may limit comparison across contexts, and sharing of indicators and lessons learned. Underlying the general preference expressed for mixed methods was an effort to balance the methodological needs of a specific evaluation (which might favor qualitative methods) with the needs of some stakeholders (who might require quantitative measures.)

Although presented as separate chapters, focused on distinct outcome areas, it may be worthwhile to consider how recommendations for organizational performance assessment may relate to evaluation of networks, partnerships and coalitions made up of organizations; or the ways in which policy influence sometimes works through organizations, networks and coalitions; or evaluation of capacity development might apply in each of these organizations or partnerships. Common challenges to evaluation in developing countries were also raised by the chapter authors, including complexity and unpredictability, and the lack of sustained resources to support ongoing monitoring and evaluation.

The editors and funder assert that the "purpose of this volume is to spark a conversation, not provide an authoritative, prescriptive statement of how to best conduct development evaluations" (p. 230). The range of experience of the authors, variety of development outcome areas addressed, and the numerous examples and specific suggestions they provided offer substantial guidance for the design, implementation and use of evaluation in international development. Multiple authors reinforced the critical importance of knowledge sharing and facilitating the use of evaluation processes and findings. This implies that effective international development evaluators must

have strong communication skills, and sufficient time to build relationships and understanding of the social and political contexts, frequently in a culture that is not their own. Many of the frameworks, models and approaches presented have originated in developed countries, and are often implemented by evaluators who come from, or have been educated in developed countries, despite their extensive experience living and working in developing countries. The growth in international evaluation associations offers hope that the voices of local evaluators will enrich our understanding of how these ideas do, and do not apply in their contexts. Ofir and Kumar sum up the need for collaboration in international development evaluation, "We need to draw from the most useful and credible international theories in development and evaluation, and build where appropriate on indigenous knowledge and value systems, experiences and wisdom" (p. 23)."

References

Pawson, R. (2006). *Evidence-based policy: A realist perspective.* Thousand Oaks, CA: Sage.

Pawson, R. & Tilley, N. (1997). *Realistic evaluation.* London: Sage.

Weiss, C. H. (1977). *Using social research in public policy making.* Lexington, MA: Lexington Books, DC Heath.

Weiss, C. H. (1982). Knowledge utilization in decision making: Reflections on the terms of the discussion. *Research in Sociology of Education and Socialization, 3, 21-27.*

About the Reviewer

Linda A. Pursely
Director of Assessment and Institutional Research
Lesley University
lpursley@lesley.edu
Linda A. Pursley, Ph.D. is currently the Director of Assessment and Institutional Research at Lesley University in Cambridge, MA. Her interest in international development was fueled by more that ten years of living and working in Africa, including assignments in Lesotho, Ethiopia, Botswana, and Zambia. She completed her doctoral studies at Cornell in Program Evaluation and Planning.