



東京大学
THE UNIVERSITY OF TOKYO



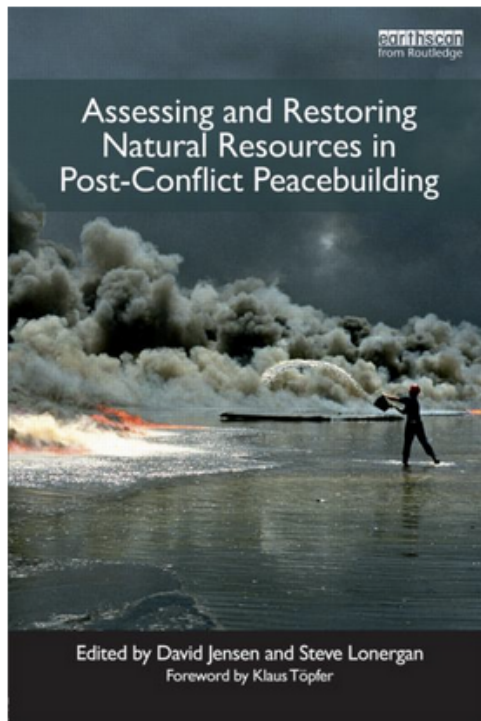
McGill



PRIO

This chapter first appeared in *Assessing and Restoring Natural Resources in Post-Conflict Peacebuilding*, edited by D. Jensen and S. Lonergan. It is one of 6 edited books on Post-Conflict Peacebuilding and Natural Resource Management (for more information, see www.environmentalpeacebuilding.org). The full book can be ordered from Routledge at <http://www.routledge.com/books/details/9781849712347/>.

© 2012. Environmental Law Institute and United Nations Environment Programme.



Evaluating post-conflict assistance

Suppiramaniam Nanthikesan and Juha I. Uitto^a

^a *United Nations Development Programme (UNDP)*

Online publication date: May 2013

Suggested citation: S. Nanthikesan, J. Uitto. 2012. Evaluating post-conflict assistance. In *Assessing and Restoring Natural Resources in Post-Conflict Peacebuilding*, ed. D. Jensen and S. Lonergan. London: Earthscan.

Terms of use: This chapter may be used free of charge for educational and non-commercial purposes. The views expressed herein are those of the author(s) only, and do not necessarily represent those of the sponsoring organizations.

Evaluating post-conflict assistance

Suppiramaniam Nanthikesan and Juha I. Uitto

In post-conflict and post-crisis settings in which people's livelihoods are at stake and situations remain fluid, there is an urgency to ensure not only that emergency responses and development interventions are on track but that they do not further exacerbate the problems. There is a real risk that external interventions in post-crisis situations could worsen inequalities that may have been the original cause of the crisis, or weaken the unifying ties among conflicting communities. Certain interventions may not be appropriate for the social, cultural, or economic situation prevailing in the area. Evaluation is a tool to ensure that an intervention is cognizant of such factors. At the same time, the evaluation must be timely.

Many evaluations are driven by the interests of the agencies and donors who wish to assess the effectiveness and efficiency of their interventions. In those cases, accountability is directed upward to the agency and its funders, frequently the taxpayers in the donor country. But what about the people the interventions are intended to benefit? Downward accountability to beneficiaries is equally if not more important. In the evaluation of post-conflict assistance, the focus on downward accountability has been weak.

This chapter explores some of the evaluation challenges particular to post-conflict situations, the importance of natural resource management to recovering from conflict, and the key role that evaluation can play in shaping effective interventions in these areas. It begins by defining evaluation and describing its role in accountability, learning, and program improvement. It examines conceptual and practical challenges to evaluating post-conflict interventions, and proceeds by discussing the limitations of the quantitative approach to evaluation, an approach currently emphasized by donors. The chapter continues with an examination of alternative approaches to effective evaluation in post-conflict situations. It discusses how evaluation can promote downward accountability and help identify unintended consequences of interventions. The chapter argues for taking a comprehensive approach to evaluation, going beyond individual interventions. To make this case, lessons for evaluating the role of natural resource management in moving from recovery to sustainable development are highlighted based on evaluations conducted.

Suppiramaniam Nanthikesan is monitoring and evaluation advisor for the Regional Bureau for Africa of the United Nations Development Programme (UNDP). Juha I. Uitto is deputy director of UNDP's Evaluation Office.

THE ROLE OF EVALUATION

Evaluation is a powerful instrument for strengthening development effectiveness. Its role ranges from summative assessments of what was achieved and measuring the program's achievements against its objectives, to formative assessments, in which the lessons learned are used to improve current and future programs. But there is no real dichotomy between accountability and learning, for both approaches require credible and verifiable evidence of the performance of the evaluand (subject of the evaluation), whether it is a project, program, strategy, policy, or organization.

There are several definitions of evaluation, but their gist is usually the same. The United Nations Development Programme (UNDP), for example, defines *evaluation* as “a rigorous and independent assessment of either completed or ongoing activities to determine the extent to which they are achieving stated objectives and contributing to decision making” (UNDP 2009, 8).

Monitoring and evaluation are often mentioned interchangeably as oversight functions. They share the aim of providing information to managers, stakeholders, and claim holders that can help inform decisions, improve performance, achieve planned results, and hold duty bearers to account.¹ However, monitoring and evaluation have important differences. Monitoring is essentially a management function, carried out by those who create and run a program, while evaluations are done independently to provide an objective assessment of whether a development initiative is on track. Such independence can be achieved either by engaging an external entity to design and carry out the evaluation or through internal mechanisms, such as an independent evaluation unit within an agency.² Evaluations are also more rigorous than monitoring in their procedures, design, and methodology, and go beyond mere description of what was achieved to an analysis of what worked and why. Most importantly, evaluations are able to bring out not only the intended results but also the unintended consequences of a development initiative—an issue that is especially important when determining whether the claim holders, or beneficiaries of a program, actually receive the benefits of an intervention.

The standards adopted by the United Nations Evaluation Group, a professional network of heads of evaluation in the United Nations system, contain principles such as the intention to use the evaluation findings; impartiality, including methodological rigor and absence of bias in the evaluation process;

¹ In this chapter, *claim holders* refers to the beneficiaries of an intervention; *duty bearers* refers to donors or authorities who have the responsibility to ensure the rights of claim holders; and *stakeholders* refers to all those who have a stake in the development intervention—primarily claim holders and duty bearers.

² The degree of such independence varies among agencies. In UNDP and the World Bank, for example, the evaluation units report directly to the respective governing bodies, which also approve their program of work and budget for evaluation. The program office whose project is being evaluated has no control or influence over how the evaluation is carried out or its results.

independence from management, to avoid undue influence and conflict of interest; and transparency and consultation with major stakeholders (UNEG 2005).

The practice of evaluation contains a wide variety of approaches. While the focus is frequently on programs or projects, it is often important to take a thematic or sector-wide approach to evaluation. This is particularly the case in post-conflict situations, where focusing on a single intervention may not be adequate. In virtually all cases there are a number of interventions consisting of a set of projects that interact and are intended to contribute to the larger outcome of conflict resolution and peacebuilding. These interventions may be nationally initiated or external, involving domestic and foreign actors, both bilateral and multilateral organizations, civil society, and others. For programmatic evaluations to be fully meaningful, they should be conducted in a manner that captures the whole picture. This is especially the case with natural resources, given their interconnected and crosscutting nature.

Timely feedback to management is critical in post-conflict settings, in which situations often evolve rapidly. Approaches have been developed for evaluating interventions in real time in order to support adaptive management. In such approaches it is important to maintain adequate distance from the program implementers, lest these evaluations slip into routine monitoring of activities and outputs.

THE CHALLENGES TO EVALUATING POST-CONFLICT INTERVENTIONS

Post-conflict evaluations face all of the challenges of routine development evaluations (that is, those not conducted in a post-conflict situation) as well as their own particular challenges. The challenges in post-conflict evaluations are both conceptual and practical.

One conceptual challenge is related to the ideal of national ownership of evaluations and the issues surrounding the legitimacy of the state. Development programming and evaluations enshrine the principle of national ownership as embodied in the Paris Declaration on Aid Effectiveness, endorsed in 2005 by over one hundred countries and agencies.³ That means that a national government assumes responsibility for an intervention and exerts effective leadership toward

³ The Paris Declaration on Aid Effectiveness defines *ownership* this way:

Partner countries commit to:

- Exercise leadership in developing and implementing their national development strategies through broad consultative processes.
- Translate these national development strategies into prioritised results-oriented operational programmes as expressed in medium-term expenditure frameworks and annual budgets (Indicator 1).
- Take the lead in co-ordinating aid at all levels in conjunction with other development resources in dialogue with donors and encouraging the participation of civil society and the private sector (OECD 2005, 3).

its success. However, because the state is itself almost always a party to the conflict, there is a concern as to whether the state can and should “own” a post-conflict evaluation. A post-conflict government may, in fact, have an interest in undermining programs that might benefit its former enemies and may therefore attempt to influence the findings of the evaluation.

Second, multiple factors affect the capacity of parties to a conflict to reconcile or to amplify their differences. The interconnected and dynamic nature of these factors poses significant challenges to fully understanding the root causes of conflict and assessing changes.

A third conceptual challenge arises from the deep fragmentation of society associated with the conflict. Peacebuilding efforts recognize this fragmentation and work to address it. But what may appear to one group as a successful or effective intervention may be seen as counterproductive by another. In this context, claims of achievements and statements about what works and why are likely to be contested. Thus the design of evaluations, particularly methods of data validation, faces significant challenges.

There are practical challenges as well. Conflicts weaken or destroy existing institutional capacity to gather data. Countries facing an urgent need to rebuild and a scarcity of resources may not be able to devote the manpower or money to create a framework for gathering evidence for evaluations. Even when data exist, their credibility and objectivity cannot be taken for granted. Thus there are significant challenges to obtaining baseline information or monitoring progress toward intended results.

Limitations of quantitative approaches

The current donor emphasis on quantitative approaches,⁴ such as results frameworks with indicators and logical frameworks,⁵ may well be premature and may crowd out efforts to achieve in-depth understanding of the assumptions and explanations surrounding conflicts (OECD/DAC 2007). Rigorous alternative evaluation approaches exist and should be used.

Many donors, facing pressure to report to their constituencies on the performance of their investments, are seeking to demonstrate tangible results and

⁴ Strictly speaking, qualitative and quantitative approaches refer to the types of data collected and not the categories of design (Weiss 1998). However, this chapter adopts the common usage whereby quantitative and qualitative approaches refer to type of designs. Quantitative approaches involve data that can be subject to statistical analysis and reporting, and these approaches focus on the size of effects and the significance of statistical relationships. Qualitative approaches rely primarily on data collected through nonstandardized methods (for instance, through semistructured or unstructured interviews) and do not have to be amenable to statistical analysis.

⁵ A results framework outlines the development hypothesis of how the program objectives are to be achieved. It presents the causal linkages between the goal, strategic objectives, and specific program outcomes, as well as the underlying assumptions.

are thus increasingly favoring quantitative approaches, in particular impact evaluations, which attempt to attribute changes in conditions to specific interventions, rather than just assessing what happened.⁶ Donor agencies frequently want to be able to identify the effects produced by their projects—often in order to be able to report such results to their funders or the taxpayers in the donor country. However, it is often impossible to isolate the effects of one particular intervention. This is especially true in complex post-conflict situations where conditions change rapidly and are influenced by multiple actors. The same can also be said for natural resource management interventions.

Impact evaluations also attempt to identify the counterfactual—what would have happened in the absence of a given intervention. They tend to emphasize experimental designs using quantitative methods derived from experimental sciences, such as randomized control trials—for example, comparing a village that received the intervention with one that did not.

Impact evaluations have shortcomings that limit their use in evaluating post-crisis responses. First, they can only be credibly used for fairly simple individual interventions that have clear and measurable objectives (Picciotto 2007). For example, in evaluating a program to vaccinate children against a specific disease, the occurrence of that disease among vaccinated and unvaccinated children could be measured, and the difference attributed to the campaign. Even in such a case, the impact evaluation would focus only on the dimension of immediate effects of the vaccination and might not look at possible unintended consequences in the community if, for instance, the campaign promoted forced vaccinations. Although in this case, the primary objective of vaccinating children would be achieved, the approach might undermine community buy-in and the long-term sustainability of the effort.

Second, real development is about not only the end result but also the means by which it is achieved. A key question for evaluations guided by human development principles is: Do the means enhance human agency and freedom? In other words, they evaluate not only the results but also the processes by which the results are achieved. Experimental (and quasi-experimental) methods, such as randomized trials, are ill-suited to handle this aspect. Furthermore, establishing the counterfactual (what would have occurred in the absence of the intervention) is rarely feasible in real-life post-crisis situations and would involve ethical challenges—if there was evidence that an intervention was indeed working, would it be ethical to deprive a control group of its benefits?

A more significant challenge comes from the issue of unintended consequences, which is an essential element of downward accountability. That is, were the intended benefits of an intervention associated with an unforeseen negative effect? However, tracking unintended consequences is by definition outside

⁶ Groups such as the Washington-based Center for Global Development and the relatively new International Initiative for Impact Evaluation (known as 3ie), based in Cairo, have been established with donor support to promote impact evaluation.

the results framework, and hence cannot be captured by experimental or quasi-experimental methods.

For these reasons, impact evaluation methods do not fit well into post-conflict and post-crisis situations, which are often messy and require complex interventions and management that can adapt to the unexpected.

In this context, rigorously assessing and identifying which programs are making a difference and which are merely feel-good exercises without real results has become a challenge. Finding a middle ground between quantitative approaches (which use experimental and quasi-experimental methods) and qualitative approaches (which rely on ethnographic research) is a familiar challenge to those who have been following the debates on development evaluations (Chen 1990).

Extending this analysis to monitoring and evaluation of peacebuilding and conflict-resolution efforts, Reina Neufeldt has presented the debate in terms of two strands of thinking among practitioners (Neufeldt 2007). *Frameworkers*, Neufeldt writes, believe that peacebuilding efforts can be interpreted through linear causal chains laid out in logical frameworks and thus can be monitored and evaluated quantitatively, using indicators that measure the achievement of predetermined activities and outputs and their contribution to higher-order objectives and goals. *Circlers*, on the other hand, take a more flexible and responsive approach, emphasizing the importance of context and the uniqueness of interventions and the communities in which they take place. They argue that events in conflict environments are unpredictable, depend on organic community processes in which causality is not clear, and thus require a qualitative approach. These two approaches reflect significant differences in understanding “how the world operates, how we interact within that world, how we can determine the impact of our programming interventions upon that world, the role of different worldviews and the purpose of determining impact” (Neufeldt 2007, 16).

EFFECTIVE POST-CONFLICT EVALUATION APPROACHES

In a post-conflict situation, the most effective evaluation approach is shaped by the nature of the conflict, the objectives and scale of the evaluation, and the resources available to carry it out. Clearly, the end of a long-standing, low-intensity conflict (for instance, Peru in the 2000s) and of a relatively brief but devastating conflict (Rwanda in the mid-1990s) present different constraints on the nature and breadth of the evaluation. Depending on the circumstances, it may be necessary to evaluate (1) specific projects or programs, (2) the performance of a particular sector (for example, health), or (3) all efforts related to a theme (for example, livelihoods). Occasionally it is necessary to take stock of the entire external assistance effort in response to a crisis (such as the Rwandan genocide).

These three types of evaluation—project, sector-wide, and thematic—share some fundamental features, such as rigorous methods, but involve different instruments and different techniques for collecting and validating data. Sector-wide

assessments often require joint evaluations involving multiple actors and often have considerable resources at their disposal. This in turn permits an elaborate and rigorous design that could include time- and resource-consuming instruments such as broad-based claim holder surveys. An evaluation of a small project, on the other hand, usually involves a much smaller budget and effort. But the smaller scale rules out certain instruments—for instance, without a substantial budget it would not be possible to conduct a survey of all claim holders. When timely evidence is needed on whether to reorient the program, real-time evaluations are necessary.

Rigorous evaluations: A theory-based approach

In the face of the challenges outlined previously, qualitative approaches to post-conflict evaluation are often more feasible than quantitative approaches. It is now well established that qualitative approaches can be rigorous. Since the 1980s the theory-based approach has been advanced as a way to conduct rigorous performance assessments that avoid the pitfalls of qualitative approaches (Birckmayer and Weiss 2000). This approach is based on a program theory—“a plausible and sensible model of how a program is supposed to work” (Bickman 1987, 5). The theory explains how the planned intervention will lead to the desired social and developmental changes and points out the risks and assumptions involved in every step. It also provides the basis for evaluations to assess progress toward results. When quantitative measures are not readily available, as is the case in many post-conflict situations, theory-driven approaches are the most viable option to ensure rigor.

Programs and projects are often initiated without a clear picture of how specific interventions will lead to the desired end result. Many post-conflict projects and programs substitute experience and intuition for theory. Even when interventions start with a theory, the dynamics of a post-conflict situation may leave the theory outdated by the time the intervention is evaluated. The role of evaluation is to scrutinize whether the theory holds and whether the activities, which are promoted in good faith, actually improve the lot of those affected by the crisis. If there have been unintended consequences, the theory needs to be revised to account for these consequences. Consequently, evaluations may need to reconstruct a theory, based on an understanding of the intervention and on consultations with the designers and implementers of the intervention.

Because relapse into conflict is a risk in most post-conflict situations, a viable theory must recognize the opportunities to bring conflicting communities together as well as the threats that could exacerbate their divisions. Constructing such a theory requires a full understanding of the root and proximate causes of the conflict, local history, traditions, institutions, power relations, and the economies of the conflicting parties. A well-constructed theory needs inputs from a range of specialists, including anthropologists, historians, economists, political theorists, and civil-society activists.

Conflict signifies a fundamental difference in the way actors interpret the world. Even in peacetime, there have been situations that required development evaluations to devise an alternative theory (Carvalho and White 2004). It may be the case that no single theory would ever be agreeable to all stakeholders. Techniques such as multi-criteria decision aid can help to gauge claim holders' values and to assess alternative strategies for program intervention (Vaessen 2006).⁷

In a theory-based approach, the program theory allows the construction of a results chain—a description of how inputs and activities lead to desired outputs and outcomes. It is then possible to construct context-appropriate indicators for the outputs and outcomes on a case-by-case basis.

Despite the advantages of a theory-based approach, there is still considerable pressure to develop and use universal indicators when planning and evaluating peacebuilding efforts. In this regard, it is worth quoting at length from an important 2007 paper on this subject by the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development:

There is no single proved methodology for preventing violence and building peace. This reality led many . . . to be concerned by recent donor emphasis on establishing standard (or universal) conflict prevention and peacebuilding indicators, specifying detailed logframe analyses of intended activities, and linking evaluations to funding decisions. Many note that they now see evaluations as existing “only” or “primarily” to meet a donor requirement.

Given the reality of what we do not know, conflict prevention and peacebuilding evaluations in the coming years should be directed toward gathering evidence and learning from it, and on testing and challenging commonly held theories and assumptions about peace and conflict, rather than on establishing fixed universal indicators of peace/conflict. Clarity on indicators (and whether or not they can be generalised in a useful way) may emerge in the process, but the focus and approach at this time should avoid over-specification of anticipated indicators as benchmarks for evaluation. Upcoming conflict prevention and peacebuilding evaluations should focus on gathering experience and analysing it cumulatively and comparatively across contexts, to improve our collective learning (OECD/DAC 2007, 14).

The progress toward the intended results thus identified can be evaluated using the DAC evaluation criteria modified and amended as necessary. Possible criteria for peacebuilding efforts include relevance and appropriateness, effectiveness, efficiency, impact, sustainability, linkages, coverage, and consistency with conflict-prevention and peacebuilding values. Achieving rigor is not feasible without

⁷ *Multi-criteria decision making* generally refers to techniques that explicitly consider multiple criteria in evaluating possible courses of action or decisions. Such decisions may have to be taken in situations that involve conflicting criteria, multiple stakeholder groups, or a lack of information (Mendoza and Martins 2006).

meaningful involvement of local experts in developing the theory of change and designing the evaluation. Involvement of local experts is equally necessary in implementing the evaluations—stakeholder mapping as well as collecting, interpreting, and validating data. The evaluation team must have the expertise in the conflict prevention and peacebuilding field, as well as the ability to interpret local customs and traditions and to fully understand the local interlocutors.

Joint evaluations

Post-conflict program support often involves multiple themes and sectors. For instance, a needs assessment has to be conducted; resources mobilized; efforts of diverse actors coordinated; relevant capacities strengthened; and links between relief, recovery, and development established. Relevant capacities of duty bearers include the ability to govern and implement law enforcement, public administration, and service delivery, while relevant capacities for claim holders include the ability to access and utilize services and to hold duty bearers accountable.

Joint evaluations are the best method to get a full picture of the peacebuilding dynamics—not only of the effects of individual efforts but also of the interactions among the multiple efforts. Joint evaluations tend to have greater objectivity and legitimacy and make it easier to capture attribution. They also are capable of strengthening downward accountability. Joint evaluations are also more effective advocacy tools to convince policy makers or program managers to address the findings and recommendations. Joint evaluations addressing the broader set of interventions are not perceived to advance any single actor's perspective. These advantages often outweigh the disadvantages associated with joint evaluations, which include higher transaction costs and complexity.

One possible method for a joint evaluation is to conduct parallel evaluations on different themes and then synthesize all evaluations to produce a single report. This approach has been followed in many thematic evaluations by UNDP and was used by the Tsunami Evaluation Coalition discussed later in this chapter (Telford, Cosgrave, and Houghton 2006). The latter evaluation was organized in five parallel thematic areas: needs assessment, coordination, impact on local and national capacities, resource mobilization, and links between relief and development. Each thematic evaluation was conducted jointly but led by one agency. A dedicated team member was assigned from the beginning to ensure that a synthesis of all thematic studies would be feasible at the end. To do so, it was necessary to ensure that evaluation questions reflected this need and used uniform standards for data collection, analysis, and reporting. A synthesis report was in fact successfully prepared.

Real-time evaluations

A real-time evaluation is conducted at the early stages of a program and focuses on the effectiveness and efficiency of implementation, rather than the ultimate results. It provides timely evidence to address ongoing programming needs in

response to the dynamic context. Procedurally, real-time evaluations can be conducted also as joint evaluations. The evaluation is conducted while a program is still being implemented, and the results are fed back to the program for immediate use. Approaches to real-time evaluation have been developed by entities that operate in crisis situations, such as the Active Learning Network for Accountability and Performance in Humanitarian Action. Agencies interviewed for a United Nations Children's Fund desk review of real-time evaluations generally found the rapid feedback valuable (Sandison 2003). Another review of real-time evaluation practice also emphasized its timeliness and usefulness to program staff (Herson and Mitchell 2005).

The risk with a real-time evaluation, however, is that it can blur the distinction between monitoring and evaluation. In such a case, a real-time evaluation (like monitoring) only focuses on program performance and delivery, rather than constantly evaluating whether the program strategy continues to make sense. Therefore, it is important to ensure the independence of the evaluation function to enable it to constantly question the assumptions behind the intervention's logic and to verify the results with the claim holders.

Real-time evaluation is just one type of rapid evaluation and assessment.⁸ A common challenge for all such methods is finding a balance between speed and trustworthiness (McNall and Foster-Fishman 2007). In order to be truly useful, real-time evaluation must be removed from program management and retain its independence, credibility, and rigor.

DOWNWARD ACCOUNTABILITY AND UNINTENDED CONSEQUENCES

In an ideal world, the accountability to the duty bearers and funders of an intervention (upward accountability) would mean the same thing as accountability to the claim holders (downward accountability), and the methodologies of evaluations commissioned by the duty bearers would be no different from those commissioned by or on behalf of claim holders. In that world, both types of evaluation would assess the full range of consequences (positive and negative, intended and unintended) that interventions had for claim holders. These would be goal-free evaluations (Scriven and Patton 1972).

However, reality is often different, by omission if not by design. Evaluations are designed to assess progress toward intended results, or goals. Results, and reporting practices based on them, guide evaluations in this direction. Evaluations commissioned by donors focus mainly on the results that are of interest to them.

⁸ Other techniques include participatory rural appraisal, which is often used to assess claim holders' perceptions of their conditions and the programs intended to benefit them; rapid evaluation methods, developed by the World Health Organization to assess the quality of health-care services; and rapid (ethnographic) assessments, associated with action research, which aim to assess local conditions and improve the design of interventions.

Claim holders, on the other hand, are concerned with any consequences, intended or unintended, that affect their lives.

Thus the key difference between upward accountability and downward accountability in the real world is how they treat unintended results. Downward accountability does not distinguish unintended results from intended results, while upward accountability generally ignores them. This might not be an issue if unintended results were minimal. However, in complex post-conflict humanitarian and recovery interventions, the unintended consequences of aid can dwarf the intended consequences (Anderson 1999; Rieff 2002; Kennedy 2004; IFRC 2005). Although evidence clearly points to the need for evaluations to go beyond assessing intended results, actual evaluation practice consistently fails to take unintended consequences seriously (Harvey et al. 2010). It seems clear that the methodologies available now are not fully equipped to capture and assess unintended consequences. Moreover, the pursuit of methodological rigor inhibits the assessment of the extent to which unintended consequences affect the claim holders.

Evaluations recognize the need to distinguish between unintended consequences that are avoidable (occur because of errors in programming) and those that are a consequence of the complex and unpredictable nature of the post-conflict situation (Anderson 1999). It is feasible for evaluations to address unintended results (Vaux 2001; Telford, Cosgrave, and Houghton 2006).

One example is the recent joint evaluation conducted by the Tsunami Evaluation Coalition, consisting of forty-two international agencies, to assess the impact of international assistance to countries affected by the 2004 Indian Ocean tsunami (Telford, Cosgrave, and Houghton 2006). In that evaluation, two of the four case studies involved regions also affected by conflict at the time of the tsunami (Sri Lanka and the Aceh region in Indonesia). The Tsunami Evaluation Coalition addressed downward accountability through four steps.

1. To assess the performance of disaster response, the evaluation set the benchmarks using international agreements and standards—the *Code of Conduct for the International Red Cross and Red Crescent Movement and Non-Governmental Organisations (NGOs) in Disaster Relief* (IFRC 1994); the Sphere Project’s *Humanitarian Charter and Minimum Standards in Disaster Response* (Sphere Project 2004); *Principles and Good Practice of Humanitarian Donorship* (Good Humanitarian Donorship 2003); and the UN’s Guiding Principles on International Displacement (UN 1998)—rather than those of the individual agencies conducting the evaluation.
2. In delineating the evaluand, which was the effect of assistance on local and national capacities, it explicitly included the ability to hold duty bearers to account as a dimension of these capacities.⁹

⁹ The evaluation defined *capacity* as the “interconnected set of skills and abilities to access services and programmes, to influence and set policies and longer term recovery/reconstruction agendas, and . . . to hold duty-bearers at all levels accountable” (Scheper et al. 2006, 17).

400 Assessing and restoring natural resources in post-conflict peacebuilding

3. The methodology identified the vulnerable and marginalized populations (taking into consideration gender, age, social status, physical disability, and geographic region) in every case study to facilitate the assessment of changes in the capacities of these groups.
4. Since downward accountability requires participatory approaches to evaluation that gauge the perceptions of claim holders, the evaluation included broad-based claim holder surveys and exit stakeholder consultations with affected communities (represented by community-based organizations) to validate its findings.

Qualitative methods were the only possible source of evidence for this evaluation. Lessons from evaluative experience also point to the need to design evaluations that clearly recognize the power relations between the duty bearers and the claim holders, as well as among all stakeholders.

Evaluations conducted by individual agencies are ill equipped to fully address downward accountability issues. To fully understand and assess the consequences of international support, sector-wide joint evaluations are seen as necessary. Evaluations conducted jointly by all key actors in a given sector will help pin down the overall contribution of international support. This became apparent in the Tsunami Evaluation Coalition study, in which claim holders could not differentiate among the different United Nations agencies and tended to regard them all as simply the United Nations (Telford, Cosgrave, and Houghton 2006). Similarly, from their perspective, very often the differences between other donor agencies and international organizations were blurred.

The tsunami, and the international assistance that followed, strengthened the efforts underway for the resolution of the conflict in Aceh.¹⁰ In Sri Lanka, the results were the opposite. Though the initial response of the population was overwhelmingly to ignore ethnic boundaries and reach out to the affected, subsequent events led to the breakdown of the fragile peace agreement. Even though aid was not the major factor leading to this breakdown, the way the national government chose to disburse international assistance became a contentious issue. This result was a consequence of a complex web of interactions between the demography of the relief and recovery needs and the political upheavals in the country.

To ensure recovery and the ability to “build back better,” it is critical that assistance programs recognize the different social groups, the social position of each, and the relationships among them. Evaluation must address whether the response is contributing to building bridges among divided social groups or merely responding to pressure to disburse funds as quickly as possible. Three of the four case studies undertaken by the Tsunami Evaluation Coalition found that aid was disbursed disproportionately to areas that were easily served by transportation, rather than based on need, and that the old and the disabled were often

¹⁰ For another perspective on post-tsunami Aceh, see Renner (2013).

excluded from benefits because they were poorly informed about them. A recovery support that clearly recognizes groups at risk and an evaluation designed to assess the differentiated progress of these different groups are vital to speedy recovery.

A number of broad-based claim holder surveys were conducted as part of the tsunami aid evaluation. For instance, in Sri Lanka and Indonesia, surveys involving over 1,000 affected people were conducted for one of the thematic components of the evaluations. These surveys provided an important channel for feedback from a number of claim holders who could not have been reached otherwise. For two reasons, the surveys would not have been possible if this had not been a joint evaluation. First, resources from all participating agencies were needed to make the survey viable. Second, the time and capacity requirements for the design and conduct of the surveys were only met because it was a joint exercise. Joint evaluations provide a necessary but not sufficient condition for experimenting to broaden the reach of the evaluation.

EVALUATING NATURAL RESOURCE MANAGEMENT IN POST-CONFLICT SITUATIONS: ASSESSMENTS OF DEVELOPMENT RESULTS

A comprehensive understanding of the context and situations affecting people's lives and well-being in post-conflict situations is essential. In order for any intervention to reach the intended beneficiaries and to improve the lives of claim holders, it is important to grasp the full picture of the political, economic, and social forces that impinge on them. Consequently, narrowly evaluating a particular project or intervention without giving consideration to the wider context may generate an irrelevant assessment at best and, at worst, miss significant unintended and unforeseen consequences of the intervention. This section discusses lessons from country-level evaluations in a range of conflict-affected countries conducted by the UNDP Evaluation Office. These evaluations—called assessments of development results (ADRs)—are better able to capture the full situation and analyze how specific interventions influence outputs because ADRs take a comprehensive look and apply an endogenous perspective. What may have to be sacrificed in this approach is attributing change to limited external interventions. Such attribution, especially in a complex post-conflict situation, would in any case be futile.

Many post-conflict societies and communities depend on natural resources for their livelihoods, yet as other chapters in this book have demonstrated, agriculture, forestry, and fisheries are often disrupted during conflict and in some cases intentionally destroyed. Thus natural resource management issues must be a key focus for evaluation as the post-conflict transition moves from emergency relief to long-term development. In that transition, evaluation must move beyond considering immediate outputs such as reduction in fighting, increased safety, and access to shelter and food. Evaluation can enhance understanding of people's needs and preferences around natural resources, and of the social dynamics—such

402 Assessing and restoring natural resources in post-conflict peacebuilding

as power relations between different groups or between genders—that is essential for any intervention to succeed.

Natural resources can be both a cause of conflict and a factor enabling post-conflict recovery and development. Their role in conflicts is well documented (UNEP 2009). Post-conflict strategies and interventions must address any conflict over natural resources in a fair, balanced, and realistic manner. Otherwise, they risk perpetuating the root causes of the conflict and jeopardizing the peace. Evaluations can play a key role in ensuring that this issue is addressed. Even when natural resource issues are not a source of conflict, peoples' long-term ability to support themselves often depends on their access to natural resources such as water and land. In post-conflict and post-disaster situations, displaced people may actually increase the pressure on resources in a way that reduces sustainability. Therefore, evaluation must focus on how the program manages access to natural resources and resolves related disputes. Management of the environment and natural resources should not be an afterthought during evaluation.

As independent evaluations, ADRs assess UNDP's contributions to development results at the country level and provide lessons for strategies in the future. The evaluations take a longer perspective than is usual for evaluations of individual interventions, normally analyzing UNDP's work in the country over a period of seven to ten years. In post-conflict countries, the perspective may be somewhat shorter, depending on whether UNDP remained present in the country during the conflict or whether it returned only after hostilities ended. In either case, the evaluations look at the full range of projects and nonproject activities (such as policy dialogue and advocacy) that UNDP has been involved in. The purpose of these evaluations is to contribute to UNDP's accountability toward the organization's executive board, as well as national stakeholders and partners, to serve as a means of quality assurance for UNDP's interventions at the country level, and to contribute to learning at the country, regional, and corporate levels. While the evaluations focus on UNDP's performance and results, these are presented against the in-country situation and development trends in the country. It is not satisfactory for the projects to reach their objectives, operating effectively and efficiently, if there is no real improvement in the development situation in the country and in the lives of the people. The country's development context and status thus become important benchmarks against which UNDP's strategic positioning and results are mapped.

On the other hand, what ultimately happens in the country depends primarily on the national actors and a wide range of other partners. It is most often difficult, if not impossible, to attribute the final development results to UNDP or any other single actor. Apart from other actors and organizations, the results depend on a number of external factors, notably the overall political, economic, and security trends. Similarly, the ways that UNDP's interventions affect change may be indirect and complex, triggering actions by various actors in the country that, combined with other actions, may lead to the desired outcomes. For instance, if UNDP supports the organization of elections in a post-conflict situation, assessing

whether the elections resulted in democratization and sustainable development at the societal level in the long term would go far beyond evaluating UNDP's performance. If UNDP's thrust is to advocate and provide policy advice to national authorities for inclusive development that reflects the needs of vulnerable populations, it is again impossible to attribute the eventual reduction in vulnerability of, for instance, female farmers to the policy dialogue UNDP had with the government. However, it is possible to identify the contributions of UNDP to processes and whether these processes led to desired results or whether there were unforeseen factors or consequences that reduced or even reversed the benefits.

The ADRs use mixed methods to collect and analyze much data and information on the national development situation and trends, as well as on UNDP's strategies and interventions. Data collection and analysis include: review of project and program documents, monitoring and review reports, and project-level evaluations; identification and interviews with a wide range of stakeholder groups; and field visits to observe the situation on the ground and to interview claim holders. It is important to spend adequate time scoping the evaluation and mapping the stakeholders before embarking on the evaluation. When identifying the stakeholders and persons to be interviewed, it is important to include different categories, such as government officials, political parties (or opposing sides in a conflict), civil society representatives, academics, journalists, international development partners, and key groups that have a stake in the process. It is also important to reach out to groups that are not involved in the interventions to gauge whether the interventions have had noticeable effects—either positive or negative—beyond the direct stakeholders. A key analytical tool used to enhance the validity of the findings is triangulation, referring to a deliberate effort to confirm information from multiple data sources, using multiple methods and repeated observations over time, and analyzing findings against multiple hypotheses (UNDP EO 2011).

Taking a long-term and comprehensive perspective to analyzing UNDP's support to two conflict-affected countries in the African Great Lakes region (Rwanda and Uganda), ADRs were able to assess the relevance, effectiveness, efficiency, and sustainability of the organization's strategies and interventions. Although the individual projects were mostly found to be operating as expected, the evaluations pointed out certain shortcomings that reduced the overall results in both countries.

Almost two decades later, Rwanda is still feeling the aftermath of the 1994 genocide. Particularly vulnerable groups include widows and orphans (in 2006, almost a quarter of Rwandan households were headed by women and 0.7 percent by children), recently returned refugees, and resettled internally displaced persons (UNDP EO 2008). Similarly, families of those detained as suspects of crimes related to the genocide are vulnerable. In a country where agriculture employs some 80 percent of the population, management of land and water is critical (UNEP 2011). In the Land of a Thousand Hills, as Rwanda is often called, it is estimated that half of the country's farmland suffers from moderate to severe

erosion (UNEP 2011). Due to its high rural population density (with 350 people per square kilometer, Rwanda is the most densely populated country in Africa) and continued population growth of 2.5 percent per year, agriculture has been pushed to steep slopes and to seasonally flooded valley bottoms that are ecologically fragile (UNEP 2011; UNDP EO 2008). During the genocide, erosion-control structures were abandoned, and they remain neglected.

Environmental management has received little attention from the government, despite the key role sustainable agriculture could play in recovery from the conflict. Looking beyond the immediate aftermath of the conflict, an ADR—covering the period from 2000 to 2006—highlighted the need for addressing issues related to sustainable management of natural resources, including neglected land management works. The evaluation found that UNDP did not adequately prioritize its assistance to the government to address this development gap and ensure the effective incorporation of environment and sustainable development into the economic development and poverty reduction strategy (UNDP EO 2008). The longer time window and comprehensive view taken in the evaluation allowed for the identification of such gaps, like the critical linkages between poverty and environment and the pressure on arable land that remains a potential source of conflict (Diamond 2005; Boudreaux 2009).

In neighboring Uganda, economic development has been one of the fastest in sub-Saharan Africa, with sustained growth rates averaging 7.8 percent since 2000. In the north, however, two decades of conflict between the government and the Lord's Resistance Army and other rebel groups took a toll on economic and social development; about 1.8 million people were internally displaced, and thousands were killed. Although the hostilities officially ended in 2006, the region still lags severely behind the rest of the country in social and economic development. Reducing this disparity is a major challenge for the government (UNDP EO 2009b).

With up to 80 percent of the population engaged in agriculture and a rapidly growing population (3.3 percent per year), sustainable use and management of the environment is key to human and economic development. As part of its early recovery and human security interventions in the northern parts of the country, UNDP has promoted livelihood projects for internally displaced persons. However, an ADR discovered that while these projects, some of which were directly related to agriculture and food security, were seen as positive, the results were mixed because of their short-term nature, implementation delays, and lack of community consultation in their design (UNDP EO 2009b).

Both evaluations reinforce the conclusion that paying attention to natural resource management and livelihoods is essential in a post-conflict situation in which most people depend on agriculture for their sustenance. They also showed the importance of integrating such considerations in policies and long-term strategies, beyond individual projects, in order to achieve sustainable benefits for the conflict-affected populations. This empirical analysis provides further evidence to the potential role of natural resources, not only as sources of conflict, but as an area around which

cooperation can be built in the Great Lakes region, provided that they are properly factored into post-conflict development visions (Kameri-Mbote 2007; UNEP 2009).

An ADR conducted in Afghanistan provided different, albeit equally critical, lessons for working on local development in a conflict situation. International attention in Afghanistan has focused understandably on insecurity and the continued conflict with the Taliban. But the evaluation showed that exclusive focus on ending conflict in insecure areas can undermine the transition to development.

The United Nations is present in Afghanistan under a Security Council mandate and operates as an integrated mission. Assistance to Afghanistan by the United Nations and the international community more broadly has focused on dealing with the immediate issues pertaining to the conflict and post-conflict recovery. Like most donors, UNDP has concentrated its efforts in Afghanistan on the provinces with the most security problems. An ADR conducted in 2009 suggested that this approach may have failed to create incentives to reduce armed conflict—as peace and security, where they have been established, have not brought significant social and economic benefits. Despite its expertise in and experience with promoting community-based development programs and concepts in many countries, UNDP as part of the integrated mission also focused on post-conflict recovery and only later became active in promoting economic development and livelihoods. Such activities are crucial, however, as the lack of jobs and income opportunities make people more likely to join insurgent activities (UNDP EO 2009a).

This echoes the conclusion from a more general evaluation of UNDP support to conflict-affected countries (UNDP EO 2006). These evaluations recommended that UNDP advocate strongly for the international community to drop the incremental, phased approach to post-conflict assistance in favor of one that immediately begins to build the capacity of the institutions necessary for lasting peace and development. This should give priority early on to issues such as reintegration of refugees and internally displaced persons and support for sustainable livelihoods. In countries that are highly dependent on agriculture, support for natural resource management is crucial.

CONCLUSIONS

Evaluation can determine whether post-conflict peacebuilding objectives are being met, and can also provide lessons for how to improve performance. In order for evaluations to provide reliable feedback, they need to incorporate a number of considerations.

Evaluators must understand the dynamics that either help or hamper people's efforts to improve their lives after conflict ends. To this end, they need to have a full grasp of the nuances of the local context—including a comprehensive understanding of the social, political, and economic causes that fueled the crisis, and of existing local capacities with which recovery efforts need to engage. They must particularly keep in mind inequalities, power relations, gender dimensions, social exclusion, and vulnerabilities that exist in the communities. The various

links to natural resource ownership, control, access, and management should also be considered.

Experience shows that the international community tends to repeat its own mistakes in responding to post-conflict and post-crisis situations because situations that in fact are quite variant are seen as interchangeable and amenable to standardized solutions. But national, political, social, economic, and ethnic issues, among others, as well as the sources of conflict, vary significantly from one situation to the next. Interventions do not take place in a vacuum; there are always other factors taking place. This variability creates challenges to the meaningful utilization of preexisting frameworks of analysis in post-conflict countries.

In post-conflict evaluations there are no easy solutions involving tick boxes or simple indicators. The evaluations should provide both upward accountability, to the funders of the program interventions, and downward accountability, to those the program is intended to benefit. Surveys, focus groups, and interviews with key informants all play a role, but they must be carefully crafted so as not to bias the findings in favor of particular groups or interests. Another important element is triangulation, or validating the findings from interviews through other means, such as document review, statistical data on actual trends, and direct observation.

Evaluations should go beyond intended results and dig deeper to look at the full spectrum of consequences, both intended and unintended, of development initiatives. Any changes in the well-being of claim holders are affected not only by the intended effects but also by the unintended consequences. In post-conflict situations, unintended consequences could be significant compared to the intended ones. The failure of most post-conflict evaluations to assess intended and unintended results on an equal footing has resulted in these evaluations not reflecting downward accountability and raises questions about their rigor and reliability.

To capture the full spectrum of consequences, evaluation approaches to recovery efforts must be based on a thorough understanding of the experiences of the local populations as collected through interviews and surveys. In addition, assessment of sector-wide efforts should be preferred over individual interventions. Effort must be made to document evidence accumulating from post-conflict evaluations to build a body of knowledge that will help test and refine assumptions and hypotheses about providing post-conflict support. At the same time, evaluation can also help improve our collective understanding regarding the role of natural resource management in both conflict and effective peacebuilding.

REFERENCES

- Anderson, M. B. 1999. *Do no harm: How aid can support peace—or war*. London: Lynne Rienner.
- Bickman, L. 1987. The functions of a programme theory. In *Using programme theory in evaluation*, ed. L. Bickman. San Francisco, CA: Jossey-Bass.
- Birckmayer, J. D., and C. H. Weiss. 2000. Theory-based evaluation in practice: What do we learn? *Evaluation Review* 24 (4): 407–431.

- Boudreaux, K. 2009. Land conflict and genocide in Rwanda. *Electronic Journal of Sustainable Development* 1 (3): 85–95.
- Carvalho, S., and H. White. 2004. Theory-based evaluation: The case of social funds. *American Journal of Evaluation* 25 (2): 141–160.
- Chen, H.-T. 1990. *Theory-driven evaluations*. Thousand Oaks, CA: Sage Publications.
- Diamond, J. 2005. *Collapse: How societies choose to fail or succeed*. New York: Viking.
- Good Humanitarian Donorship. 2003. *Principles and good practice of humanitarian donorship*. Stockholm, Sweden. www.goodhumanitarianandonorship.org/Libraries/Ireland_Doc_Manager/EN-23-Principles-and-Good-Practice-of-Humanitarian-Donorship.sflb.ashx.
- Harvey, P., A. Stoddard, A. Harmer, and G. Taylor. 2010. *The state of the humanitarian system: Assessing performance and progress; A pilot study*. With V. DiDomenico and L. Brander. London: Active Learning Network for Accountability and Performance in Humanitarian Action. www.alnap.org/pool/files/alnap-sohs-final.pdf.
- Herson, M., and J. Mitchell. 2005. Real-time evaluation: Where does its value lie? *Humanitarian Exchange Magazine* 32:43–45.
- IFRC (International Federation of Red Cross and Red Crescent Societies). 1994. *Code of conduct for the International Red Cross and Red Crescent movement and non-governmental organisations (NGOs) in disaster relief*. Geneva, Switzerland.
- . 2005. *World disasters report 2005: Focus on information in disasters*. Bloomfield, CT: Kumarian Press.
- Kameri-Mbote, P. 2007. *Conflict and cooperation: Making the case for environmental pathways to peacebuilding in the Great Lakes region*. Environmental Change and Security Program Report 12. Washington, D.C.: Woodrow Wilson Center / United States Agency for International Development. www.wilsoncenter.org/sites/default/files/ECSP_report_12.pdf.
- Kennedy, D. 2004. *The dark side of virtue: Reassessing international humanitarianism*. Princeton, NJ: Princeton University Press.
- McNall, M., and P. G. Foster-Fishman. 2007. Methods of rapid evaluation, assessment, and appraisal. *American Journal of Evaluation* 28 (2): 151–168.
- Mendoza, G. A., and H. Martins. 2006. Multi-criteria decision analysis in natural resource management: A critical review of methods and new modelling paradigms. *Forest Ecology and Management* 230 (1–3): 1–22.
- Neufeldt, R. C. 2007. “Frameworkers” and “circlers”—Exploring assumptions in peace and conflict impact assessment. Berlin, Germany: Berghof Research Center.
- OECD (Organisation for Economic Co-operation and Development). 2005. Paris declaration on aid effectiveness. Paris. www.unrol.org/files/34428351.pdf.
- OECD/DAC (Organisation for Economic Co-operation and Development, Development Assistance Committee). 2007. Encouraging effective evaluation of conflict prevention and peacebuilding activities: Towards DAC guidance. *OECD Journal on Development* 8 (3): 7–106.
- Picciotto, R. 2007. The new environment for development evaluation. *American Journal of Evaluation* 28 (4): 509–521.
- Renner, M. 2013. Post-tsunami Aceh: Successful peacemaking, uncertain peacebuilding. In *Livelihoods, natural resources, and post-conflict peacebuilding*, ed. H. Young and L. Goldman. London: Earthscan.
- Rieff, D. 2002. *A Bed for the night: Humanitarianism in crisis*. New York: Simon and Schuster.

408 Assessing and restoring natural resources in post-conflict peacebuilding

- Sandison, P. 2003. Desk review of real-time evaluation experience. Evaluation working paper. New York: United Nations Children's Fund.
- Scheper, E., A. Parakrama, S. Patel, and T. Vaux. 2006. *Impact of the tsunami response on local and national capacities*. London: Tsunami Evaluation Coalition.
- Scriven, M., and M. Q. Patton. 1972. Pros and cons about goal-free evaluation. *Evaluation Comment: The Journal of Educational Evaluation* 3 (4): 1–7.
- Sphere Project. 2004. *Humanitarian charter and minimum standards in disaster response*. Geneva, Switzerland. <http://ocw.jhsph.edu/courses/refugeehealthcare/PDFs/SphereProjectHandbook.pdf>.
- Telford, J., J. Cosgrave, and R. Houghton. 2006. *Joint evaluation of the international response to the Indian Ocean tsunami: Synthesis report*. London: Tsunami Evaluation Coalition.
- UN (United Nations). 1998. Guiding principles on internal displacement. E/CN.4/1998/53/Add.2. www.unhcr.org/refworld/docid/3d4f95e11.html.
- UNDP (United Nations Development Programme). 2009. *Handbook on planning, monitoring and evaluating for development results*. New York: web.undp.org/evaluation/handbook/documents/english/pme-handbook.pdf.
- UNDP EO (United Nations Development Programme, Evaluation Office). 2006. *Evaluation of UNDP assistance to conflict-affected countries*. New York. <http://web.undp.org/evaluation/documents/thematic/conflict/ConflictEvaluation2006.pdf>.
- . 2008. *Assessment of development results: Rwanda; Evaluation of UNDP contribution*. New York. www.oecd.org/dataoecd/61/0/41105593.pdf.
- . 2009a. *Assessment of development results: Islamic Republic of Afghanistan; Evaluation of UNDP contribution*. New York. www.undp.org.af/publications/KeyDocuments/ADR_Afghanistan.pdf.
- . 2009b. *Assessment of development results: Uganda; Evaluation of UNDP contribution*. New York. http://web.undp.org/evaluation/documents/ADR/ADR_Reports/Uganda/Uganda_ADR_2009.pdf.
- . 2011. *ADR method manual*. New York.
- UNEG (United Nations Evaluation Group). 2005. *Norms for evaluation in the UN system*. New York. www.uneval.org/papersandpubs/documentdetail.jsp?doc_id=21.
- UNEP (United Nations Environment Programme). 2009. *From conflict to peacebuilding: The role of natural resources and the environment*. Nairobi, Kenya. www.unep.org/pdf/pcdmb_policy_01.pdf.
- . 2011. *Rwanda: From post-conflict to environmentally sustainable development*. Nairobi, Kenya. http://postconflict.unep.ch/publications/UNEP_Rwanda.pdf.
- Vaessen, J. 2006. Programme theory evaluation, multicriteria decision aid and stakeholder values: A methodological framework. *Evaluation* 12 (4): 397–417.
- Vaux, T. 2001. *Independent evaluation: The DEC response to the earthquake in Gujarat*. London: Disasters Emergency Committee.
- Weiss, C. 1998. *Evaluation*. 2nd ed. Upper Saddle River, NJ: Prentice-Hall.