

20 Evaluation in conflict resolution and peacebuilding

Esra Çuhadar Gürkaynak, Bruce Dayton, and Thania Paffenholz

Introduction

Conducting evaluation is rarely a favorite activity for those engaged in conflict resolution and peacebuilding work (hereafter CR/PB). It takes time, consumes scarce resources, requires a relatively high degree of expertise, and can result in evaluation results that are already self-evident or do not capture the nuances of conflict transformation work. Yet there are good reasons to bring evaluation to the forefront of the field of CR/PB. First, evaluation is an essential instrument for monitoring and improving upon existing initiatives. Without solid evaluation, practitioners would lack the ability to understand “what went wrong,” and scholars would lack the ability to build a body of theory about the causes of and remedies to social conflicts. Second, as the number of non-governmental and international organizations involved in peacebuilding activities increases, so too does the call for greater accountability on the part of these organizations. Finally, evaluation is an almost universal obligation when it comes to fulfilling the terms of conflict management grants given by public and private donors.

In recent years scholars have answered the call for evaluation in CR/PB work in a number of ways. Some have offered new approaches to conceptualizing the meaning of success in conflict resolution interventions, reconciliation initiatives, and other peacebuilding efforts (Mitchell 1993; d’Estree *et al.* 2001; Ross 2004; Rouhana 2000). Others have highlighted particular case studies in an effort to demonstrate the conditions and contexts that lead to “more-or-less successful” outcomes (Douma and Klem 2004; Lieberfeld 2002). Still others have outlined specific questions or frameworks that could be usefully applied to CR/PB activities (Anderson and Olson 2003; Church and Shouldice 2002, 2003; NPI-Africa 2002; Paffenholz and Reychler 2005, 2007). Despite this impressive body of work, many practitioners in the CR/PB field remain skeptical about the overall merits and usability of traditional evaluation to their work. Some have argued that the complexity of CR/PB work makes outcome and impact evaluation nearly impossible to conduct, and that traditional program evaluation tools are incapable of measuring the kinds of intangible changes that occur during conflict resolution initiatives (Anderson and Olson 2003).

The purpose of this chapter is to examine this debate, to examine the difficulties and possibilities of applying program and policy evaluation frameworks, methodologies and tools to CR/PB work, and to illustrate how traditional program and policy evaluation can be effectively used and/or modified and adapted to CR/PB assessment frameworks. The chapter proceeds in three sections: a general overview of the concept of evaluation as it is applied to the field of CR/PB; a review of the state of the art in evaluation frameworks and methodologies now being used in the field; and a discussion of various challenges currently confronting CR/PB practitioners.

Understanding the concept of evaluation

Before discussing evaluation in CR/PB, we take a look at the concept of evaluation in general. A widely agreed definition of evaluation is provided by Rossi *et al.* (1999: 16): “an evaluation is a systematic assessment of policies, programs, or institutions with respect to their conception and implementation as well as the impact and utilization of their results.” In scientific research, policy evaluation is an established discipline, which is concerned with the effects of public policies (Rossi *et al.* 1999; Bussmann 1997). Evaluation is also well established in the fields of development and humanitarian action.

An evaluation can have a number of different and simultaneous objectives including: reviewing and judging current status in order to improve interventions; controlling processes and procedures for purposes of accountability; assessing and documenting what has been achieved; identifying lessons learned for use in future interventions. Evaluations can occur at any time during the implementation process. Usually, however, evaluations are conducted in the middle of an intervention (mid-term review) or at the end (*ex post* evaluation). Evaluations of long-term, complex programs or institutions can also take place periodically, starting shortly after the beginning of the implementation process, in order to further direct the intervention and allow for any necessary adjustments in course.

Essentially, there are two types of evaluations, formative and summative. Formative evaluations seek insight into ways to improve the intervention in question and focus on process, whereas summative evaluations assess and judge the intervention’s quality and success in meeting its objectives and focus on outcome.

Over the past few years, participatory and utilization-focused evaluation processes have become widespread. These processes involve primary stakeholders in the evaluation process, and emphasize the importance of stakeholders’ ability to use the results for future improvement (Patton 1997). This approach reflects an orientation to evaluation that defines success in terms of how well the results can be used by those involved in the intervention. The participation of primary stakeholders in the evaluation process not only optimizes the use and acceptance of the results, but also contributes to joint institutional learning.

Evaluation research and practice draws on a set of criteria that are used in order to direct assessment. Each evaluation criterion includes a number of questions and issues to be explored and addressed by the evaluator(s). These questions can be answered through the evaluators’ application of different evaluation methods. Often there exists a set of “standard criteria,” so called on account of their use by most stakeholders in evaluation. For example, the OECD criteria and the European Commission’s additional criteria are well known and widely used for the evaluation of development programs. Criteria to be used for a particular evaluation are decided upon when those involved in the planning of an evaluation stipulate which issues should be evaluated. In general, the criteria of “relevance” (the significance of the intervention for its set goals and donor policies), “effectiveness” (changes an intervention has achieved with respect to its immediate environment), “impact” (changes an intervention has achieved with respect to its larger context), and “efficiency” (cost-effectiveness) are used in evaluating all types of interventions. Additional criteria used for development evaluations include “sustainability,” “coordination,” and “coherence,” and the criteria of “coverage,” “protection,” and “participation” are often applied in humanitarian evaluations. “Participation” is a criterion that assesses whether beneficiaries of interventions have been sufficiently involved in the implementation process.

The following prerequisites are necessary before one can adequately conduct an evaluation (Paffenholz and Reychler 2007: 42).

- 1 Clear and measurable objectives must be defined for the intervention. If the objectives are too vague, an accurate assessment will not be possible.
- 2 A baseline study must be conducted prior to the intervention so that a before/after comparison can be made as part of the evaluation. According to the OECD/DAC Glossary (OECD/DAC 2002), a baseline study is an analysis describing the situation prior to an intervention, against which progress can be assessed.
- 3 Results chains and indicators are best to assess the results of the intervention as a means to understand and verify the underlying hypotheses of change due to an intervention.

What methods will be used for an evaluation depends first and foremost upon the objective of the evaluation in question and the data to be analyzed. It also depends on the level of scientific rigor that is to be applied to an evaluation, and the size of the available budget. Whereas some methods (e.g. large-*N* studies and experiments) are better suited for generalization, testing, and “explaining,” others do not lead the way to generalization, but are useful for understanding special cases and their idiosyncrasies. Evaluation studies using multiple methods reveal more effective and triangulated results.¹

The range of methodological approaches used for evaluation can be grouped along two key dimensions based on a conventional distinction in social science research.² The first dimension is whether the evaluator is seeking to analyze the effects of an intervention (i.e. the changes accomplished through the intervention) through *etic* measures (as defined by the outsider evaluator) or *emic* measures (as defined by the insiders to the program or policy). There are examples of evaluation studies adopting either approach that we further elaborate in the “state of the art” section. Druckman (2005) provides a useful typology of methodologies along these two dimensions. *Etic* and *emic* approaches can be both qualitative and quantitative. For example, whereas the structured-focused comparison method is *etic* and qualitative, experiments, surveys, or aggregate case comparisons are *etic* and quantitative. *Etic* approaches to evaluation research often use surveys, experiments, and focused comparisons (or aggregate case comparisons), whereas *emic* approaches use interviews, focus groups, and participant observation in order to explore the “insider” view.

In sum, evaluation is a complex methodological endeavor, especially if the goal is impact assessment. An intervention’s impact is determined by examining the larger changes initiated by the intervention within the general context, changes that often occur only after a longer time has passed. To attribute these changes to the intervention in question is often difficult as there may be many other reasons why certain changes have occurred. Such attribution problems are common in impact assessment and are referred to as the “attribution gap.”

Evaluation in peacebuilding

The issue of evaluation has only recently entered the field of CR/PB. This delay has to do with the evolution of the field as such. In the mid-1990s we witnessed a rapid increase in peacebuilding activities by a variety of actors ranging from international and regional organization to academic institutions, foundations, civil society groups, social movements, business groups, and the media. Consequently, a decade later the field felt sufficiently

matured to reflect on its own learning. Parallel to these learning efforts, a debate about professionalization in CR/PB has recently emerged. It is in the context of these different developments that the evaluation debate gained momentum in CR/PB.

Evaluation in peacebuilding is needed to respond to a set of interrelated needs and purposes. First, the results of the evaluation of single CR/PB interventions provide the intervening actors and stakeholders with information on how to improve the intervention. Second, evaluations of CR/PB interventions help strengthen the accountability of the intervening organization vis-à-vis its respective constituencies and donors. Fourth, evaluations can support a culture of reflection and learning. Fifth, evaluations enhance the general learning in the CR/PB field. Finally, evaluations help practitioners and scholars refine their theories about the causes and dynamics of conflict, thus enabling them to refine their approaches to CR/PB.

Resistance and current trends

Despite these benefits, peacebuilding actors have often resisted evaluation for a number of reasons. Some in the field fear that the essential goals and values of peacebuilding can simply not be measured – in line with Albert Einstein’s quote, “not everything that can be counted counts, and not everything that counts can be counted.”

In the same vein goes the argument that peacebuilding evaluations also differ from other evaluations in terms of their specific context of armed conflicts and peace processes, which are incredibly complex social and political phenomena. Peacebuilding processes are also intensely vulnerable, making them difficult to assess in the short run; ultimately, only sustainable peacebuilding counts as success. Thus, it is necessary to evaluate whether current interventions are on the right road to contribute to sustainable peacebuilding.

Peace researchers also argue that the current debate over evaluation misses the mark as serious research on the effects of peacebuilding interventions is more important than conducting evaluations. This argument is often based on a misperception of evaluation and is due, in part, to the blurry distinction between evaluation and research. Many wrongly equate evaluation with poorly conducted evaluations reliant on rushed processes. Indeed, there is a danger that many evaluations are implemented in a quick and rushed manner; however, this does not discredit the concept of evaluation as such.

In sum, peacebuilding interventions have some special characteristics that need to be taken into account for evaluation purposes. How these specifics can be addressed and incorporated into evaluation designs has been the subject of a research and practitioner discourse that started in the beginning of the new millennium. In the next section, we discuss these recent developments and the current state of the art.

State of the art: overview of ideas, concepts and frameworks

Although generic frameworks that can be used for evaluating peacebuilding initiatives are rare, there has been an increasing amount of research on the evaluation of peacebuilding recently. The existing scholarly work in this area can be divided into four categories, none of which is mutually exclusive: first, “lessons learned” studies, mainly commissioned by donor agencies; second, research-oriented case studies (single or comparative) that provide an in-depth study of a single or multiple peacebuilding initiatives; third, studies suggesting key questions for and/or reflection about evaluation in peacebuilding; fourth, overall frameworks and methodologies suggested for the evaluation of peacebuilding initiatives.

In the following paragraphs we give examples of the existing work on evaluation in each of these categories. We then present a typology of several evaluation frameworks that have contributed to research on evaluation of peacebuilding. Our goal in presenting such a typology is to better compare these studies, approaches, and frameworks and understand where they stand in terms of the program and policy evaluation literature. This overview is by no means a fully exhaustive one, but rather is based on a sample of evaluation studies.

Lessons learned studies

During recent years, a number of studies have been published to document lessons learned from peacebuilding initiatives. In some of these studies, the focus is on a specific country or region in which the donor organization has funded peacebuilding work. In others, different programs are analyzed from an organizational learning perspective. The motivation behind these publications is often the requirement to report lessons learned by the donor agencies, or accumulation of knowledge, or keeping archival and organizational records. Regardless of the motivation, these kinds of studies have made an important contribution to the development of CR/PB evaluation. They not only serve as a guideline for other practitioners, but also contribute to the formation of a database for further research.

A recent example of this type of research is a study prepared by the Swedish International Development Cooperation Agency (SIDA) (Sorensen *et al.* 2000). This report synthesizes lessons learned in peacebuilding projects drawn from a number of other “lessons learned” studies conducted by several donor agencies, such as USAID and UNDP. Other prominent examples of “lessons learned” evaluation studies have been the Mott Foundation study for their peace and conflict resolution projects in former communist countries (see Mayer *et al.* 1999); the German Development Cooperation study concerning projects in more than fifty countries (see Paffenholz and Brede 2004); the Joint Utstein Study (Smith 2003); and the Reflecting on Peace Practices (RPP) report (Anderson and Olson 2003).

In sum, evaluation studies undertaken in a “lessons learned” format often recommend practical suggestions that can be adopted by other practitioners and donor agencies. Many of the “lessons” concern how the organization can be more effective in designing and conducting future peacebuilding initiatives. Although making an important contribution to practice, such evaluation studies have often neglected to carry out a systematic assessment of participants’ and locals’ (emic) view of the initiative, focusing instead on the (etic) views of organizers.

Research-oriented case studies

A second type of evaluation in peacebuilding is research-oriented case studies, often focusing on outcomes and impacts or their process contributions, either of a single country/initiative or of several initiatives in a comparative perspective. This type of evaluation is, in general, research oriented and designed to build or test theory. Case studies use a variety of methods, and can be either emic or etic in their approach to research. For example, whereas Anderson and Olson (2003) adopted an emic approach to develop criteria for what is effective in peace practices based on the reflections and conceptualizations of practitioners, Maoz (2004) used criteria found in the literature on intergroup relations and attitude change to evaluate the effects of the Israeli–Palestinian coexistence initiatives. Furthermore, whereas some of these studies are descriptive and focus on the idiosyncratic aspects of a particular conflict or initiative,³ others prefer comparative case study designs

(Çuhadar 2004; Fisher 2005; Lund 2002; Susskind *et al.* 1999); and still others combine surveys and experiments to test the applicability of certain theoretical frameworks in specific peacebuilding cases (Atieh *et al.* 2005; Maoz 2004; Rosen 2006; Ohanyan and Lewis 2005; Cuhadar-Gurkaynak and Genc 2006).

Finally, some research-oriented case studies are interested in impact and outcome at the micro (Maoz 2004; Malhotra and Liyanage 2005) and/or macro levels (Lieberfeld 2002), whereas others concentrate on documenting and describing the process from a theoretical point of view (Kelman 2005). It should also be mentioned that most of the research-oriented studies focus on the evaluation of dialogue, peace education initiatives, and to some extent CR training.

Key questions/reflection studies

The third type of evaluation study in peacebuilding offers key questions for and/or reflection about evaluation. Even though some of these studies are referred to as overall frameworks, they are different from overall frameworks, although they can be used in combination with them. Rather than measurement indicators or standard evaluation criteria, these studies provide a list of questions that could serve as a guideline for those who plan and design CR/PB initiatives or those who evaluate them. Examples of such evaluation studies are Church and Shouldice (2002), the RPP initiative, and Fast and Neufeld's work (2005).

A prominent example of an evaluation study of this type is the Church and Shouldice study (2002: 26–7). In this study, the authors provide guiding questions structured around three themes: goals and assumptions behind the CR initiatives; process accountability with regard to the operationalization of the peacebuilding initiative; and the range of results from the initiative in the short and long terms. The authors then formulate a useful set of questions for each theme. For instance, the category “goals and assumptions” lists questions that help the practitioners to reflect on the appropriateness of the intervention (e.g. intervention strategy, activities), theoretical analysis (e.g. the theory of change adopted by the practitioner), and strategic review (e.g. whether the organization is doing what it says it is doing).

Another prominent study conducted in this tradition is the guidebook prepared by the RPP project (Anderson and Olson 2003; CDA 2004). It is hard to categorize the RPP study into only one of the four categories we discuss here, since it has elements that reflect all of the four categories. It can be considered a research-oriented comparative case study; a study conducted with a “lessons learned” objective; a guideline that provides questions for reflection; and even an overall framework to some extent because it offers several effectiveness criteria and a general methodology that can be used especially for formative evaluation. Still, we mention it in this section because the core of the framework (at least in its current form) offers guiding questions derived from the comparative case research on peace NGOs. These guiding questions urge the peace practitioners to consider key elements of effectiveness in their program design and they include whether the change generated by the initiative was fast enough, sustained, large enough, and linked to other levels (Anderson and Olson 2003: 16). The framework delineates a planning process based on the notion of “theory of change,” which practitioners can use when they are planning peace programs. In this sense, it can be used as a formative evaluation tool even though the research part of the study was conducted as an *ex post* summative evaluation.

Evaluation studies conducted in this style are useful especially in encouraging

practitioners to engage in reflective practice. In addition, they can be used in combination with overall evaluation frameworks that we discuss in the following paragraphs. However, they are not adequate per se since they do not suggest indicators or methodologies for practitioners or evaluators.

Overall frameworks and methodologies studies

A fourth type of evaluation study formulates general frameworks. These studies often aim at introducing methodologies and/or criteria/indicators for the assessment of peacebuilding initiatives. Some of these frameworks focus on particular peacebuilding initiatives (d'Estree *et al.* 2001); others develop generic frameworks, such as action evaluation, that can be used as an overall methodology regardless of the type of peacebuilding activity (Ross and Rothman 1999; Paffenholz and Reychler 2007). For instance, a 2005 thematic issue of the *Journal of Peacebuilding and Development* (2:2) focuses on evaluation and gathers together a collection of articles that include both general and more specific approaches to peacebuilding evaluation, as well as a few integrated approaches to development/peacebuilding evaluations. Here we also find proposals on how to merge the essential values of the CR/PB field with strategy-oriented processes and methods. A number of US-based NGOs have published guidelines for monitoring and evaluating NGO conflict transformation programs (Church and Rogers 2006).

Some prominent examples in this category are a recent book by Paffenholz and Reychler (2007), which includes guidelines for the evaluation process; a methodological reference based on the "Aid for Peace" framework and evaluation criteria as used in policy, development, and humanitarian work; and many tools and examples from field testing. Another evaluation framework has been offered by Rothman and Friedman (2002) and Ross and Rothman (1999) in their "action evaluation" approach. This framework applies an organizational action research paradigm to the conflict resolution interventions. It focuses on goal setting and goal fulfillment at the organizational level as important components that influence the effectiveness of conflict resolution initiatives. Ross and Rothman (1999) established a computer-based interactive process that helps stakeholders in the initiative to clarify their organizational goals and priorities as they implement their activities. In this sense, the framework is a planning approach that is more appropriate at early stages of project development. Although the framework is a significant contribution for organizations at the goal-setting stage, it does not address impact assessment or suggest indicators by which to measure impacts.

Two evaluation frameworks by d'Estree and her colleagues suggested a combination of evaluation methodologies and criteria and indicators for assessment. The frameworks can be used for both formative and summative evaluation. One of these frameworks concerns evaluation of environmental conflict resolution (ECR) initiatives (d'Estree and Colby 2000), whereas the other is concerned specifically with the evaluation of interactive conflict resolution (ICR) workshops (d'Estree *et al.* 2001). D'Estree and Colby's (2000) framework on ECR initiatives developed categories and criteria that are appropriate for such initiatives as well as a guidebook that can be used by evaluators. Assessment criteria were developed for each of the following categories: outcome reached, process quality, outcome quality, relationship of parties to outcome, relationship between parties, and social capital. The guidebook was intended to be used as a tool for standardized case assessment, research and evaluation strategy, organizational framework, and education. Similarly, the d'Estree *et al.* (2001) evaluation framework for ICR suggested categories and assessment criteria for the

following: changes in representation (referring to cognitive changes in the participants), changes in relations, foundations for transfer, and foundations for outcome/implementation. These categories and criteria tried to capture both micro- and macro-level changes that occur as a result of ICR workshops.

Table 20.1 lists several of the evaluation studies discussed so far and compares them in terms of the key evaluation concepts that we introduced early on in this chapter. We realize that this table includes only a sample and is not exhaustive. Our purpose is to show the variety of evaluation approaches that exist in CR/PB.

Reflections and challenges

We now review several distinct challenges that confront CR/PB practitioners as they seek to evaluate their work. These challenges include articulating a theory of change at the outset of CR/PB work, overcoming the “attribution gap,” reconciling the evaluation preferences of donors with those of “local” stakeholders, and collecting data in conflict zones. We briefly reflect on each of these challenges below.

Theories of change

All attempts to intervene in conflict situations begin with a set of assumptions about the nature of the conflict, the factors that keep it from being resolved, and the means by which it can be transformed. Conflict scholars refer to these assumptions as a “theory of change” and argue that they are foundational to the design of all CR/PB efforts. Theories of change serve as an implicit result chain that predicts how an initiative’s activities proceed through a series of steps to a desired outcome (Weiss 1998).

All evaluation specialists are in agreement that articulating a theory of change, i.e. making it explicit in the planning phase, is an essential step in helping CR/PB managers understand what they are doing, why they are doing it, and how they can determine whether or not their objectives have been achieved once the activity has been concluded. Yet this task often proves to be quite difficult. Conflict analysis is a notoriously complex undertaking full of competing propositions about the forces that cause conflicts to emerge and the means through which these forces can be transformed. As a result, theories of change are rarely articulated at the outset of CR/PB work as CM/PB practitioners struggle to connect the theories they hold to the activities they conduct. To rectify this problem, three steps are necessary.

First, practitioners should always explicitly articulate the general beliefs they hold about the conflict with which they are dealing before designing specific project activities. These beliefs could be categorized in the following way: (1) beliefs about the underlying roots/bases of the conflict being investigated (where it “comes from”); (2) assumptions about how these bases of the conflict are causally linked; (3) beliefs about the conditions under which these root causes can be transformed (in either a positive or a negative direction); and (4) beliefs about what kinds of programmatic interventions bring about what kind of transformations.

Second, CR/PB practitioners need to be able to articulate how the impact of their work will transfer from the micro to the macro level; that is, describe how the effects (outcomes and impacts) that any given initiative may have on individual participants will channel their way “up” to broader societal institutions, narratives, and practices (Anderson and Olson 2003; Çuhadar 2004; Fisher 2005; Mitchell 1993; Ross and Rothman, 1999;

Table 20.1 A typology of evaluation studies

	Objective	Timing	Type	Any standard/criteria developed?	Evaluator	Level of evaluation	Type of peace-building intervention
Action evaluation framework (Ross and Rothman 1999)	Review and judge current status to improve	Mid-term or <i>ex ante</i>	Formative, participatory	No	Insiders with the help of an outsider	Micro	Not specified
Reflecting on peace practice (Anderson and Olson 2002)	Identify lessons learned	<i>Ex post</i>	Summative, formative	Some	Outsider	Macro	All kinds of initiatives
Paffenholz evaluation framework (Paffenholz and Reyhler 2007)	Review and judge current status to improve and walk user through process and application	All stages (baseline, mid-term, and <i>ex post</i>)	Summative, formative, participatory	Yes	Insider or outsider	All levels	All kinds of initiatives
Environmental conflict resolution evaluation (d'Estree and Colby 2000)	Assess and document what has been achieved; develop criteria and indicators for evaluation	All stages (baseline, mid-term, and <i>ex post</i>)	Formative and summative	Yes	Insider or outsider	Micro and macro	Environmental conflict resolution initiatives
Peace education evaluation (Abraham Fund and IPCRI sponsored initiatives in Israel; Maoz 2000, 2004)	Assess and document what has been achieved, identify lessons learned	<i>Ex post</i>	Summative	Yes	Outsider	Micro	Peace education/coexistence initiatives
D'Estree <i>et al.</i> evaluation framework (d'Estree <i>et al.</i> 2001)	Develop criteria and indicators for evaluation	All stages (baseline, mid-term, and <i>ex post</i>)	Formative and summative	Yes	Insider or outsider	Micro, meso, and macro	Interactive conflict resolution

Rouhana 2000; Paffenholz and Reychler 2007). For instance, how do initiatives that target attitudinal or behavioral changes at the individual level get transferred to those outside the group who were not involved in the activity? How does the strengthening of local civil society institutions in post-conflict zones lead to the creation of societies that are more resistant to the re-emergence of destructive conflict? Theories of change able to lay out these transfer processes will facilitate the design of appropriate indicators of change at the *micro*, *meso*, and *macro* levels (Mitchell 1993).

Third, theories of change in CR/PB work should take into account the “stage” of the conflict that is being addressed (see Susan Allen Nan’s Chapter 27 in this volume). The impact of CR/PB initiatives is likely to vary depending on whether the initiative takes place at the emergence phase of the conflict, the escalation phase, the pinnacle of the conflict, the de-escalation phase, or the post-conflict reconciliation stage. At each of these stages, a different set of psychological and material dynamics are at play, each of which interacts with the CR/PB initiatives in different ways. Part of building a theory of change in the intervention’s planning phase, therefore, involves consciously selecting activities that one theorizes are most likely to be successful at the present stage of the conflict. So, for instance, traditional peacekeeping interventions may make more sense at an acute phase of a conflict than at the escalation phase (Fisher and Keashly 1991). Similarly, establishing truth and reconciliation commissions may make more sense at the post-conflict stage than at other stages.

Addressing the attribution gap

A second fundamental challenge of CR/PB evaluation involves finding a way to attribute the activities that were carried out by the CR/PB practitioner to the change that is observed. This so-called “attribution gap,” discussed earlier, is familiar to anyone who conducts evaluation (Rossi *et al.* 1999). Nonetheless, it is perhaps amplified in conflict settings where multiple stakeholders, with multiple interests and conflict perspectives, interact in a crisis environment. Unlike a controlled laboratory experiment, where it is easy to assess how changing one variable changes the overall environment, conflict zones are messy places where a multitude of uncontrollable external variables constantly impact the dynamics on the ground. Under these circumstances, the impact of particular CR/PB activities is difficult to see against the backdrop of large external forces playing upon the conflict. For instance, economic hardships, shifting regional political alliances, leadership transitions, or even the activities of diaspora communities thousands of miles away can directly or indirectly impact the ability of the CR/PB initiative to meet its objectives. Similarly, months of work getting different ethnic groups to sit down together for a confidence-building dialogue could be lost after only one day of violence between external members of those groups in areas far away from the dialogue. When analyzing initiative success, therefore, evaluators face the daunting challenge of needing to “control” for the impact of external variables on the success of the initiative.

Controlling for external variables is even more difficult in regions where multiple organizations are conducting different initiatives, involving different populations and activities, at the same time. How can the analyst identify macro-level impacts of one particular initiative when other initiatives, with different objects, participants and activities, are being conducted simultaneously? In many cases, the answer to these questions may be “one cannot.” At times, it may be true that conflict phenomena are so complex that we can never know with certainty that correlations between specific initiatives and an overall

reduction of conflict are anything but spurious (see Dennis Sandole's Chapter 30 in this volume). This does not, however, mean that evaluation should be abandoned. On the contrary, practitioners might instead seek to limit claims about the impacts of particular CR/PB activities to those that can be validated through carefully designed evaluative methods. They might also give more attention to conducting baseline studies prior to initiating CR/PB initiatives so that initiative impacts are rendered visible. Further, they should develop novel approaches to conceptualizing indicators of change, and donors and implementers can then join hands and commission public opinion polls able to capture these change indicators.

Reconciling the evaluation preferences of donors with those of local stakeholders

A third challenge to CR/PB evaluation concerns the importance of including local voices in the evaluation process. A report by Anderson and Olson (2003), for instance, notes that a significant difference exists between the perceived goals of donors, practitioners, and target populations when it comes to evaluation. On the one hand, the donor community has been characterized as pursuing an evaluation agenda that is driven by goals such as efficiency, timeliness, sustainability, and coherence, and based on the establishment of pre-determined and verifiable indicators of change (Hoffman 2001). On the other hand, these kinds of indicators may stand in sharp contrast to the needs and interests of the stakeholders involved in the activity. For these individuals, quantitative measures of the initiative's success and the goals of efficiency may be secondary to more immediate needs, such as ceasing hostility, providing safe havens for refugees, or delivering essential services.

Fortunately, there has been an upsurge in "user-driven" or participatory approaches across the CR/PB spectrum. These approaches emphasize the involvement of stakeholders in the construction of evaluation designs, the creation of success indicators, and the modification of the initiative's goals over the life of the activity. This tends to create a more "organic" assessment of problems and needs, and makes practitioners more accountable to the community at which the initiative is targeted (Bock 2001). Participatory approaches to CR/PB evaluation also help to overcome the challenge of devising culturally appropriate evaluation models and debunk the myth that evaluation is a Western practice that misses the nuances of the cultural variation across conflict zones.

Difficulties of data collection in conflict zones

Finally, many conflict zones are not dangerous places, but many others are. In these areas, such as Kashmir, Darfur, or the West Bank and Gaza, conducting interviews and collecting data with which to evaluate initiatives may be problematic, or impossible. In some cases CR/PB personnel may be available to answer questions and talk about initiative impacts but the stakeholders who are affected by these changes are not. Here the risk is that CR/PB practitioners might present a more optimistic picture of what has been accomplished than would the stakeholders who are affected by the intervention. If evaluators are unable or unwilling to traverse insecure territory, bringing participants/stakeholders together in a neutral location outside the immediate conflict area may be a good option for capturing critical evaluation data. Another option can be to commission data collection to local groups that might have easier access or to rely on the results of self-evaluation by involved stakeholders (Paffenholz and Reychler 2007). Depending on the quality of the data, success indicators need to be limited to those that can be assessed.

Conclusion

The call for more rigorous evaluation in the CR/PB field will only intensify over the years to come. Official donor evaluation guidelines, greater professionalization of the field, growing numbers of organizations doing CR/PB work, and calls for greater accountability are now transforming the way that CR/PB practitioners approach the assessment of their work. Adapting to this new reality, however, does not have to be unpleasant. Despite misgivings about the time, resources, and complexity of evaluation, renewed attention to evaluation has several benefits: it forces practitioners to translate their sometimes vaguely held theories of change into concrete plans for implementation; it directs attention to the attribution gap and ways to overcome it; it strengthens the connection between CR/PB theory and practice; and it encourages practitioners to find new indicators of conflict transformation that demonstrate the central role that civil society groups, educational institutions, NGOs, and other CR/PB organizations now play in conflict transformation.

The rich array of evaluation ideas, concepts, cases, and frameworks reviewed in this chapter indicate that those involved in evaluating CR/PB work have made tremendous strides over a short period of time in charting a path forward. Collectively, this work shows that practitioners now have many good choices when it comes to capturing the results of their work. It also suggests that evaluation can become a natural and automatic part of CR/PB planning, implementation, review, and adjustment. Indeed, so much CR/PB work of such variety is now being conducted in every corner of the world that it is, perhaps for the first time, now possible to outline the conditions and context that make CR/PB work more or less successful. Evaluation is the key to unlocking these secrets and to building a more professional field that is rigorous both in theory and in application.

Notes

- 1 See Dennis Sandole's Chapter 30 in this volume on "Critical Systemic Inquiry." See also Druckman and Stern (2000) and Druckman (2005). For an application see Cuhadar-Gurkaynak and Genc (2006).
- 2 See Druckman (2005) for an overview of a range of methods used in conflict resolution.
- 3 See Agha *et al.* (2003) on the Arab–Israeli conflict; Kelman 2005 on the Israeli–Palestinian conflict; Paffenholz (2005, [2003] 2006) on the Life and Peace Institute's work in Somalia; Voorhees (2002) on the Dartmouth process.

Bibliography

- Agha, H., Feldman, S., Khalidi, A., and Schiff, Z. (2003) *Track II Diplomacy: Lessons Learned from the Middle East*. Cambridge, MA: MIT Press.
- Anderson, M. and Olson, L. (2003) *Confronting War: Critical Lessons for Peace Practitioners*. Cambridge, MA: Collaborative for Development Action.
- Atieh, A., Ben-Nun, G., El Shahed, G., Taha, R., and Tulliu, S. (2005) 'Peace in the Middle East: P2P and the Israeli–Palestinian Conflict', United Nations Institute for Disarmament Research (UNIDIR) report, Geneva, January 2005.
- Bock, J. (2001) 'Towards participatory communal appraisal', *Community Development Journal*, 36(1): 46–153.
- Bussmann, W. (1997) *Einführung in die Politikevaluation (Introduction to Political Evaluation)*. Basel/Frankfurt: Helbing & Lichtenhahn.
- CDA (2004) *Reflecting on Peace Practice Project*. Cambridge, MA: Collaborative Learning Projects, CDA. Online. Available at <www.cdainc.com/rpp/docs> (accessed 3 February 2007).

- Church, C. and Rogers, M. (2006) *Designing for Results: Integrating Monitoring and Evaluation into Conflict Transformation Programs*. Washington, DC: Search for Common Ground/Alliance for Peacebuilding/US Institute for Peace.
- Church, C. and Shouldice, J. (2002) *The Evaluation of Conflict Resolution Interventions: Framing the State of Play*. Londonderry: INCORE.
- Church, C. and Shouldice, J. (2003) *The Evaluation of Conflict Resolution Interventions, Part II: Emerging Practice and Theory*. Londonderry: INCORE.
- Çuhadar, E. (2004) *Evaluating Track Two Diplomacy in Pre-Negotiation*. PhD dissertation, Syracuse University, Syracuse, NY.
- Çuhadar-Gürkaynak, E. and Genc, O. (2006) 'Evaluating peacebuilding initiatives using multiple methodologies: lessons learned from a Greek-Turkish peace education initiative', paper read at the International Conference on Education for Peace and Democracy, Antalya, Turkey, 19-23 November.
- Douma, N. and Klem, B. (2004) *Civil War and Civil Peace: A Literature Review of the Dynamics and Dilemmas of Peacebuilding through Civil Society*. Netherlands Institute of International Relations, unpublished report.
- Druckman, D. (2005) *Doing Research: Methods of Inquiry for Conflict Analysis*. Thousand Oaks, CA: Sage.
- Druckman, D. and Stern, P. (2000) 'Evaluating interventions in history: the case of international conflict resolution', *International Studies Review*, 2: 33-63.
- d'Estree, T. and Colby, B. (2000) *Guidebook for Analyzing Success in Environmental Conflict Resolution Cases*. Fairfax, VA: Institute for Conflict Analysis and Resolution, George Mason University.
- d'Estree, T. P., Fast, L. A., Weiss, J. N., and Jakobsen, M. (2001) 'Changing the debate about "success" in conflict resolution efforts', *Negotiation Journal*, 20(2): 101-13.
- Fast, L. and Neufeld, R. (2005) 'Envisioning success: building blocks for strategic and comprehensive peacebuilding impact assessment', *Journal of Peacebuilding and Development*, 2(2): 24-41.
- Fisher, R. (2005) 'Analyzing successful transfer effects in interactive conflict resolution', in R. Fisher (ed.) *Paving the Way*. Lanham, MD: Lexington Books.
- Fisher, R. and Keashly, L. (1991) 'The potential complementarity of mediation and consultation within a contingency model of third party intervention', *Journal of Peace Research*, 28(1): 28-42.
- Hoffman, M. (2001) *Peace and Conflict Impact Assessment Methodology*. Berlin: Berghof Research Center for Constructive Conflict Management.
- Kelman, H. (2005) 'Interactive problem-solving in the Israeli-Palestinian case: past contributions and present challenges', in R. Fisher (ed.) *Paving the Way*. Lanham, MD: Lexington Books.
- Lieberfeld, D. (2002) 'Evaluating the contributions of track two diplomacy to conflict termination in South Africa, 1984-90', *Journal of Peace Research*, 39(3): 355-372.
- Lund, M. (2002) 'Evaluating NGO peacebuilding initiatives in Africa: getting beyond good intentions or cynicism', paper read at the International Studies Association Annual Convention, New Orleans, 27 March.
- Malhotra, D. and Liyanage, S. (2005) 'Long-term effects of peace workshops in protracted conflicts', *Journal of Conflict Resolution*, 49(6): 908-24.
- Maoz, I. (2000) 'An experiment in peace: reconciliation aimed workshops of Jewish-Israeli and Palestinian youth', *Journal of Peace Research*, 37(6): 721-36.
- Maoz, I. (2004) 'Coexistence is in the eye of the beholder: evaluating inter-group encounter interventions between Jews and Arabs in Israel', *Journal of Social Issues*, 60(2): 437-52.
- Mayer, B., Moore, C., Wildau, S., and Ghais, S. (1999) *Reaching for Peace: Lessons Learned from Mott Foundation's Conflict Resolution Grantmaking 1989-1998*. Flint, MI: Charles Stewart Mott Foundation.
- Mitchell, C. (1993) 'Problem-solving exercises and theories of conflict resolution', in D. Sandole and H. van der Merwe (eds.) *Conflict Resolution: Theory and Practice*. Manchester: Manchester University Press.

- NPI-Africa (2002) *Strategic and Responsive Evaluation of Peacebuilding: Towards a Learning Model*. Nairobi: National Council of Churches of Kenya.
- OECD/Development Assistance Committee (2002) *The DAC Principles for the Evaluation of Development Assistance, Glossary of Key Terms in Evaluations and Results-Based Management, Evaluations and Aid Effectiveness*, Series No. 6. Paris: OECD.
- Ohanyan, A. and Lewis, J. (2005) 'Politics of peace-building: critical evaluation of interethnic contact and peace education in Georgian–Abkhaz Peace Camp, 1998–2002', *Peace and Change*, 30(1): 57–84.
- Paffenholz, T. ([2003] 2006) *Community-based Bottom-up Peacebuilding: Development of the Life & Peace Institute's Approach to Peacebuilding and Lessons Learned from the Somalia Experience, 1990–2000*. Uppsala: Life and Peace.
- Paffenholz, T. (2005) 'Comparative advantage of NGO peacebuilding: the role of the Life and Peace Institute in Somalia, 1990–2003', in O. Richmond and H. Carey (eds.) *Subcontracting Peace: NGOs and Peacebuilding in a Dangerous World*. Aldershot: Ashgate Publishers.
- Paffenholz, T. and Brede, D. (2004) *Lessons Learnt from the German Anti-Terrorism Package (ATP): Possibilities and Limits of Development Cooperation for Crisis Prevention and Peace Building in the Context of Countries at Risk from Terrorism*. Eschborn: GTZ.
- Paffenholz, T. and Reychler, L. (2005) 'Towards better policy and programme work in conflict zones: introducing the "Aid for Peace" approach', *Journal of Peacebuilding and Development*, 2(2): 6–23.
- Paffenholz, T. and Reychler, L. (2007) *Aid for Peace: A Guide to Planning and Evaluation for Conflict Zones*. Nomos: Baden-Baden.
- Patton, M. (1997) *Utilization Focused Evaluation: The New Century Text*. London: Sage Publications.
- Rosen, Y. (2006) 'Does peace education in the regions of intractable conflict can change core beliefs of youth?', paper presented at IPCRI International Conference on Peace Education, Antalya, Turkey, 19–23 November.
- Ross, M. (2004) 'Conceptualizing success in conflict resolution interventions', *Peace and Conflict Studies*, 11: 1–18.
- Ross, M. and Rothman, J. (1999) 'Issues of theory and practice in ethnic conflict management', in M. Ross and J. Rothman (eds.) *Theory and Practice in Ethnic Conflict Management: Theorizing Success and Failure*. London: Macmillan.
- Rossi, P. H., Lipsey, M. W., and Freeman, H. E. et al. (1999) *Evaluation: A Systematic Approach*. Thousand Oaks, CA: Sage Publications.
- Rothman, J. and Friedman, V. (2002) 'Action evaluation for conflict management organizations and projects', in J. Davies and E. Kaufman (eds.) *Second Track Diplomacy/Citizens' Diplomacy*. Lanham, MD: Rowman and Littlefield.
- Rouhana, N. (2000) 'Interactive conflict resolution: issues in theory, methodology, and evaluation', in P. Stern and D. Druckman (eds.) *International Conflict Resolution after the Cold War*. Washington, DC: National Academy Press.
- Smith, D. (2003) *Towards a Strategic Framework for Peacebuilding: The Synthesis Report of the Joint Utstein Study on Peacebuilding*. Oslo: PRIO.
- Sorensen, N. N., Stepputat, F., and Van Hear, N. (2000) *Assessment of Lessons Learned from SIDA Support to Conflict Management and Peace Building*. Stockholm: Swedish International Development Cooperation Agency.
- Susskind, L., McKearnon, L., and Carpenter, S. (1999) *Consensus Building Handbook*. Thousand Oaks, CA: Sage Publications.
- Voorhees, J. (2002) *Dialogue Sustained: The Multilevel Peace Process and the Dartmouth Conference*. Washington, DC: United States Institute of Peace.
- Weiss, C. (1998) *Evaluation: Methods for Studying Programs and Policies*. Upper Saddle River, NJ: Prentice Hall.