Evaluation in Conflict Resolution and Peacebuilding

Esra Çuhadar Gürkaynak, Bruce Dayton, Thania Paffenholz

INTRODUCTION

Conducting evaluation is rarely a favorite activity for those engaged in conflict resolution and peace-building 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 that 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 peace building 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 better 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 peace building efforts (Kriesberg 2002; 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 peace building and conflict resolution activities (Anderson and Olson 2003; Church and Shouldice 2002 and 2003; NPI-Africa 2002; Paffenholz and Reychler 2005 and 2007). Yet despite this impressive body of work, many practitioners in the PB/CR fields remain skeptical about the overall merits and usability of traditional evaluation to their work. Some have argued that the complexity of conflict resolution and peace-building 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 outline 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." Evaluation is used in almost all areas of life in theory and practice. In scientific research, for example, policy evaluation is an established discipline, which is concerned with the effects of public policies (Rossi et al. 1999; Bussmann 1997). Evaluation is well established in the fields of development and humanitarian action, and international donors have agreed on a set of evaluation criteria that are applied as a matter of course.

An evaluation can have a number of different and simultaneous objectives including: reviewing and judging current status in order to improve current interventions; checking and 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), depending on the objective of the 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, which can be further divided into several subtypes. 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 in evaluation practice there exist a set of 'standard criteria,' called such due to 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 the evaluation is planned, i.e., 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), 'efficiency' (cost-effectiveness) are used in evaluating all types of interventions. Additional criteria used for development evaluations include 'sustainability,' 'coordination,' and 'coherence,' while 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 (2002), a baseline study is an analysis describing the situation prior to an intervention, against which progress can be assessed or comparisons made at a later stage of the intervention.

(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. Using information provided by the baseline study, the intervention's stakeholders will develop results chains and corresponding indicators for monitoring and evaluation purposes during the planning.

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 which is to be applied to an evaluation, and the size of the available budget. While some methods (e.g., large-N studies and experiments) are better suited for generalization, testing, and 'explaining', others don't lead the way to generalization, but are useful to understand special cases and their idiosyncrasies. Thus, a combination of different social science and related methods can be used for evaluation depending on the purpose of the evaluation (i.e., measuring effects or evaluating the process goals and design) and data available. Using a 'mixed-method design' in evaluation is of great value if the time and budget of evaluation allows. Evaluation studies that are done by using multiple methods reveal more effective and triangulated results.[1]

The range of methodological approaches used for most 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 and frameworks that have been done with either approach that we further on elaborate in the 'state of the art' section in this paper. The second dimension is the degree of precision the evaluation research chooses for assessment, which is provided by the techniques used for data collection and analysis. This dimension is often referred to as the qualitative and

quantitative distinction in research methods. Druckman (2005) provides a useful typology of methodologies based on these two dimensions. 'Etic' and 'emic' approaches can be both qualitative and quantitative. For example, while 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, focused comparisons (or aggregate case comparisons), while 'emic' approaches use interviews, focus groups, and participant observation in order to explore the 'insider' view.

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

History of the Debate

The issue of evaluation has only recently entered the field of CR/PB. Although individual scholars had long conducted peace research within a variety of academic disciplines, it was not until the 1960s that conflict analysis and peace research were established as an interdisciplinary academic field. Initially the field focused mainly on inter-state conflict

and the management of conflicts resulting from the Cold War. More recently the focus shifted away from wars between states to the management and resolution of armed conflicts within states (Miall et al. 1999; Eriksson et al. 2003; Gurr and Davies 2002). Such conflicts have accounted for approximately 80-90% of all conflicts since 1989 (Eriksson et al. 2003: 594). Other fields of research have adopted a social-psychological perspective to concentrate on different forms of inter-personal and inter-group conflicts related to identity, racism and prejudice (see chapter by Byrne and Senehi in this volume). Yet, other fields of research focus on conflicts related to the war on terrorism, the effects of globalization, and North-South tensions.

In recent years the term 'peacebuilding' has become a common and overarching term used to describe the entire field (Schirch 2006). Initially, interest in peacebuilding grew slowly, but the mid-1990s witnessed a rapid increase in peacebuilding activities by a variety of actors ranging from international and regional organization, such as the UN, the EU or the African Union, 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. This debate includes discussions about professional standards for staff working in CR/PB as well as for intervention designs, implementation and impact. It is in the context of these different developments that the evaluation debate gained momentum in CR/PB.

Why is evaluation in conflict resolution and peacebuilding necessary?

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 design and target the intervention more to its respective goals. Second, evaluations of CR/PB interventions – especially those conducted in a participatory way involving owners and stakeholders of the intervention – help strengthen the accountability of the intervening organisation vis-à-vis its respective constituencies. Third, evaluations of CR/PB interventions hold intervening organizations accountable to their donors. Fourth, evaluations can support a culture of reflection and learning among the involved stakeholders. 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, transforming armed conflicts into peaceful means of managing disputes and encouraging social change, can simply not be measured because they are not technical issues. This argument is often based on a perception that evaluation involves the search for quantifiable measures to what many in the CR/PB field believe are non-quantifiable transformations 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. While other evaluations also need to analyze their respective contexts, armed conflicts are incredibly complex social and political phenomena that require special attention and in-depth analysis. 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, especially in terms of assessing outcomes and particular impact (the job of social science research). Many wrongly equate evaluation with poorly conducted evaluations reliant on quickly 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. Especially in the fields of development and humanitarian action, discussion about standard evaluation criteria and best practices is on-going. If organizations fail to exploit this available information, it is not the fault of evaluation as a concept. We have, however, found insufficient research-oriented evaluations as well as a lack of evaluation-oriented research projects, especially when it comes to impact assessment. Consequently,

peace researchers should better collaborate with donors and agencies in order to encourage research-oriented studies that complement impact assessments.

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

Generic frameworks that can be used for evaluating peace-building initiatives are rare. Yet, in the last few years there has been an increasing amount of interest and research on the evaluation of peace-building. 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 peace-building initiative, or a number of initiatives, or impact assessment of a single peace-building case using multiple methods. Third, studies suggesting key questions for and/or reflection about evaluation in peacebuilding. Fourth, overall frameworks and methodologies suggested for the evaluation of peace-building initiatives.

In the following paragraphs we give examples of the existing work on evaluation belonging to each of these four categories. We then present a typology of several evaluation frameworks that have contributed to research on evaluation of peace-building in order to facilitate a discussion of these frameworks. We formulated this typology using key components of the policy and program evaluation literature that we introduced earlier in this essay. 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 important articles and studies have been published by scholars, practitioners, and donor agencies to document lessons learned from peace-building initiatives. In some of these studies, the focus is on a specific country or region in which the donor organization has funded peace-building work. In others, the approach is to combine various experiences gathered from an analysis of different programs from an organizational learning perspective. The motivation behind these publications varies: some are written because they are required by the donor agencies; others are written to contribute to the accumulation of knowledge in this area; and still others are prepared simply for archival and organizational purposes. Regardless of the motivation, these kinds of studies have made an important contribution to the development of CR/PB evaluation. The documented experiences not only serve as a guideline for other

practitioners, but also contribute to the formation of a database for further research on such initiatives.

A recent example of this type of research can be given from an evaluation study prepared by the Swedish International Development Cooperation Agency (SIDA) (Sorensen et. al. 2000). This report synthesizes lessons learned in peace-building projects drawn from a number of other 'lessons learned' studies conducted by several donor agencies, such as USAID and UNDP. The synthesized 'lessons learned' list generated by SIDA used the ALNAP criteria (Active Learning Network on Accountability and Performance in Humanitarian Assistance): appropriateness and timeliness, coherence and co-ordination, connectedness, cost-effectiveness, coverage and impact/outcome (Sorensen et. al. 2000: 31). Another example of a 'lessons learned' evaluation study was prepared by the Mott Foundation. The report lists a number of lessons learned derived from several Mott Foundation funded peace and conflict resolution projects in former communist countries (Mayer et. al. 1999). This study was based on review of written documents, observation, questionnaires sent to staff and directors of partner organizations, interviews with partner organizations, and site visits. Still, another example of a 'lessons learned' study is the ATP-study commissioned by the German Development Cooperation that assessed a global conflict prevention and peacebuilding fund with projects in more than 50 countries. The study was based on five qualitative case studies and analysis of written documents and questionnaires sent to staff in more then 40 countries (GTZ; Paffenholz and Brede 2004).

Finally, other prominent examples of lessons learned studies include the 'Joint Utstein Study' (Smith 2003), which analyzed the peacebuilding efforts funded by several governments and the 'Reflecting on Peace Practices report (RPP)', which evaluated lessons learned from NGO peacebuilding efforts (Anderson and Olson 2003).

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 organizational learning, and how the organization can be more effective in designing and conducting future peace building initiatives. For example, there are several lessons learned regarding the necessity of successful collaboration between international partners and local staff in the projects mentioned in the Mott Foundation and SIDA reports. 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 peace-building is research-oriented case studies, often focusing on outcomes and impacts or their process contributions, either of a single country/initiative or several initiatives in a comparative perspective. This type of study is, in general, research oriented and designed to build or test theory. Case studies conducted in this fashion use a variety and a combination of methods, and can be either emic or etic

in their approach to research. For example, while 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 inter-group relations and attitude change to evaluate the effects of the Israeli-Palestinian co-existence initiatives. Furthermore, while 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 (Cuhadar 2004; Fisher 2005; Lund 2002; Susskind, Kearnon, and Carpenter 1999); and still others combine surveys and experiments to test the applicability of certain theoretical frameworks in specific peacebuilding cases (Athie 2004; 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), while 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 studies in peacebuilding offers key questions for and/or reflection about evaluation in peacebuilding. Even though some of these studies are also

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 PB/CR initiatives or those who evaluate them. Examples of such evaluation studies are Church and Shouldice (2002), the 'Reflecting on Peace Practices' initiative, which was carried out by the 'Collaborative for Development Action', and Fast and Neufeld's work (2005).

A prominent example of an evaluation study of this type is the Church and Shouldice (2002: 26-27) study. 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 formulated a useful set of questions for each theme. For instance, the category goals and assumptions list 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 frequently mentioned evaluation study conducted in this tradition is the guidebook prepared by the 'Reflecting on Peace Practice Project' (Anderson and Olson 2003; CDA Report 2004). It is hard to categorize the RPP evaluation 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. Despite the difficulty to classify this study, 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. Guiding questions 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 and in guiding them towards meaningful reflection. In addition, these types of studies 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. Often, these kinds of studies aim at introducing methodologies and/or criteria/indicators for the assessment of peace-building initiatives. Some of these frameworks focus on particular peace-building initiatives (d'Estree et.al 2000); others develop generic frameworks, such as action evaluation, that can be used as an overall methodology regardless of the type of peace-building activity (Ross and Rothman 1999; Paffenholz and Reychler 2007). For instance, a 2005 thematic issue of the Journal of *Peacebuilding and Development* (vol. 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).

Also in this category is a recent book by Paffenholz and Reychler (2007), which provides a chapter on evaluation in peacebuilding with guidelines for the conceptual, process and methodological aspects of evaluation. These include guidelines for the evaluation process (including a self-evaluation by stakeholders prior to an external evaluation); 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 frequently cited evaluation framework has been offered by Rothman and Friedman (2002) and Ross and Rothman (1999) in their 'action evaluation' approach to evaluation. This framework applies a specific organizational action research paradigm to the conflict resolution interventions. The authors focus on goal setting and goal-fulfilment at the organizational level as important components that influence the effectiveness of conflict resolution initiatives. To achieve this, 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 which is more appropriate at early stages of project development. It does not suggest any standards or indicators of success and is more concerned with operational evaluation. Although the Rothman procedure 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 prepared 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 initiatives (d'Estree and Colby 2000), while the other is concerned specifically with the evaluation of interactive conflict resolution workshops (d'Estree et al. 2001). d'Estree and Colby's (2000) framework on environmental conflict resolution (ECR) initiatives developed categories and criteria that are appropriate for ECR 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 interactive conflict resolution suggested categories and assessment criteria pertaining to each of these categories. The categories and criteria suggested by the authors tried to capture both micro and macro level changes that occur as a result of interactive conflict resolution workshops. The categories are: changes in representation (referring to cognitive changes in the participants), changes in relations, foundations for transfer, and foundations for outcome/implementation.

Table 1 below 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 a sample of these studies and is not an exhaustive one. Our purpose is to show the variety of evaluation approaches that exist in CR/PB.

Table 1 A Typology of Evaluation Studies

INSERT TABLE 1 HERE

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 war zones. We briefly reflect on each of these challenges below.

Theories of change

Regardless of the conflict being addressed, the kinds of people that are brought together to address it, or the plans and resources committed, 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 consist of: (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 (either in a positive or negative direction); and (4) beliefs about what kinds of programmatic interventions bring about what kind of transformations. This sequence of assumptions serves often 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.

Unfortunately, theories of change are rarely articulated at the outset of CR/PB work. While practitioners may have a vague sense of why they are doing what they are doing, this rarely includes a well-thought-out result chain that can be used as a basis for measurement. To make the work of articulating a theory of change easier, we suggest that practitioners articulate their theory of change prior to designing their activity as part of a good planning culture. Three suggestions may be helpful in this regard.

First, practitioners need to be able to articulate the theoretical perspective that they hold as they approach their work. 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. Ross (2004:2), for instance, has identified six different "theories of practice" prevalent in CR/PB work, each of which revolves around distinct "beliefs about the nature of social, political, and psychological reality". These different theories of practice are lenses through which different practitioners approach CR/PB. Some, for instance, emphasize structural violence and institutionalized inequalities when diagnosing conflicts and designing intervention strategies (Galtung 1969); others prefer to focus on interest-based dynamics (Fisher and Ury 1983); others on the beliefs and narratives that conflicting parties tell about each other; and still others on the unhealed emotional wounds, trauma, and victimhood psychology (Linder 2002). In their recent work, Paffenholz and Reychler (2007) propose three methodological approaches to identify a theory of change through clarifying a vision for peacebuilding with the help of: (a) using internationally agreed norms and standards, (b) using research results to measure success, and (c) developing one's own

framework, for example, with the use of action evaluation. Being aware of the lenses through which practitioners are approaching their work is a critical first step in articulating a theory of change.

Second, as with interventions of all kinds, CR/PB practitioners need to be able to articulate how the impact of their work will transfer from the micro to the macro levels; 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/Olson 2003; Cuhadar 2004; Fisher 2005; Mitchell 1993; Ross and Rothman, 1999; Rouhana 2000, Paffenholz and Reychler 2007). For instance, how do initiatives that target attitudinal or behavioural changes at the individual level get transferred to those outside of 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 28 in this volume). Conflicts are fluid social processes that may be more amenable to settlement at certain stages of their evolution than at other stages (Kriesberg 2007). Zartman (2000), for instance, has argued that the parties in a conflict need to reach a 'hurting stalemate' prior

to their being ready to seriously engage in peace processes. Similarly, negotiation specialists argue that competing parties may lack the will to negotiate so long as they believe that they can achieve their interests through confrontation alone. The point here is that 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. Being able to articulate the logic of these choices through a clearly defined theory of change is essential to good evaluation. Including the development of future conflict scenarios into planning is also an option for coping with this challenge. It is an instrument that enables peacebuilders to adapt their activities to changing conflict dynamics and can be part of a conflict analysis (Paffenholz and Reychler 2007: 19).

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. Indeed, 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. For example, we have identified nearly three dozen major citizen-based peacebuilding initiatives that were conducted between Israelis and Palestinians between the late 1990s and 2006. 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 32 in this volume). This does not, however, mean that evaluation should be abandoned. Quite 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. Peace practitioners often argue that external evaluations ignore the knowledge of local actors, despite the fact that it is these very populations that PB/CR work typically seeks to empower. 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 which is driven by goals such as efficiency, timeliness, sustainability, and coherence, and which is 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.

The good news is that standard evaluation procedure today has become much more adept at involving local actors on the evaluation team as well as in the evaluation process through participatory workshops, briefings, and de-briefings. Indeed, in recent years, 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 that is the target of the initiative (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, and briefly, many conflict zones are not dangerous places, but many others are. In these areas, such as the West Bank/Gaza, Kashmir, or Darfur, 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 which is outside of 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

Like it or not, the call for more rigorous evaluation in the CR/PB field will only intensify over the years to come. Official donor evaluation guidelines, greater professonalization 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. The good news is that adapting to this new reality 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 a 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.

Bibliography

Agha, H., et al, (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: The Collaborative for Development Action, Inc.

Atieh, A. et al. 2004 'Peace in the Middle East: P2P and the Israeli-Palestinian Conflict', United Nations Publications UNIDIR, Geneva.

Bock, J. (2001) 'Towards participatory communal appraisal,' *Community Development Journal*, 36 (1): 46-153.

Bussmann, W. (1998) *Einführung in die Politikevaluation (Introduction to Political Evaluation)*, Basel/Frankfurt: Helbing & Lichtenhahn.

Church, C. and Shouldice, J. (2002) *The Evaluation of Conflict Resolution Interventions: Framing the State of Play*, Londonderry, Northern Ireland: INCORE.

Church, C. and Shouldice, J. (2002) and (2003) *The Evaluation of Conflict Resolution Interventions, Part I and II*, Londonderry, Northern Ireland: INCORE.

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.

Collaborative Learning Projects, CDA. (2004) Reflecting on Peace Practice Project. [Online]. Collaborative Learning Projects, CDA. Available from: www.cdainc.com/rpp/docs on 3 February 2007. [cited 28 February 2007].

Çuhadar, E. (2004) *Evaluating Track Two Diplomacy in Pre-Negotiation.* Ph.D. Dissertation, Syracuse University, Syracuse, NY.

Cuhadar-Gurkaynak, 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 Democarcy, Antalya, Turkey.

d'Estree, T. and Colby, B. (2000) *Guidebook for Analyzing Success in Environmental Conflict Resolution Cases*. Fairfax, Virginia: The Institute for Conflict Analysis and Resolution, George Mason University.

d'Estree, T, et al. (2001) 'Changing the debate about "success" in conflict resolution efforts', *Negotiation Journal* 20 (2): 101-113.

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: Sage.

Druckman, D. and Stern, P. (2000) 'Evaluating interventions in history: The case of international conflict resolution', *International Studies Review*, 2: 33-63.

Eriksson M. and Wallensteen, P. (2003) 'Armed conflict, 1989 – 2003', *Journal of Peace Research*, **41** (5): 625-636.

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. and Ury. W. (1981) *Getting to Yes*, New York: Viking Penguin.

Fisher, R. (2005) 'Analyzing successful transfer effects in interactive conflict resolution', in R. Fisher (ed.) *Paving the Way*, Lanham: Lexington Books.

Fisher, R. and Keashly, L. (1991) 'The potential complimentary of mediation and consultation within a contingency model of third party intervention', *Journal of Peace Research*, 28(1): 28-42.

Galtung, J. (1969) 'Violence, peace and peace research', *Journal of Peace Research*: 167-191.

Gurr, T. and Davies, J. (2002) 'Dynamics and management of ethnopolitical conflicts', in J. Davies and E. Kaufman (eds.) *Second track diplomacy/citizens' diplomacy*, Lanham, MD: Rowman and Littlefield.

GTZ, (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, GTZ/BMZ paper series on Crisis Prevention, Conflict Transformation and Peace-Building, Bonn/Eschborn, by Thania Paffenholz with Dunja Brede.

Hoffman, M. (2001) *Peace and conflict impact assessment methodology*, 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: Lexington Books.

Kriesberg. L. (2004) 'The relevance of reconciliation actions in the breakdown of Israeli-Palestinian negotiations', *Peace and Change*, 27 (4): 546-571.

Kriesberg, L. (2007). *Constructive Conflicts*, Lanham: Rowman and Littlefield.

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.

Linder, E. (2002) 'Healing the cycles of humiliation: How to attend to the emotional aspects of unsolvable conflicts and the use of humiliation entrepreneurship', *Peace and Conflict Journal of Peace Psychology*, 8 (2): 125-138.

Lund, M., (2002) *Evaluating NGO Peacebuilding Initiatives in Africa: Getting Beyond Good Intentions or Cynicism*, Paper read at the International Studies Association Annual Convention, 27 March 2002, New Orleans.

Malhotra, D. and Liyanage, S. (2005) 'Long-term effects of peace workshops in protracted conflicts', *Journal of Conflict Resolution*, 49 (6): 908-924.

Maoz, I., (2000) ‘An experiment in peace: Reconciliation aimed workshops of Jewish-Israeli and Palestinian youth’, *Journal of Peace Research*, 37 (6): 721-736.

———. (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): 437-452.

Mayer, B., et. al. (1999) *Reaching for Peace: Lessons Learned from Mott Foundation's Conflict Resolution Grantmaking 1989-1998*. Flint, Michigan: Charles Stewart Mott Foundation.

Miall, H., Ramsbotham, O., and Woodhouse, T. (1999) *Contemporary Conflict Resolution*, Cambridge: Polity Press.

Mitchell, C. (1993) ‘Problem-solving Exercises and Theories of Conflict Resolution’, in D. Sandole and H. Merwe (eds.) *Conflict Resolution: Theory and Practice* Manchester University Press.

NPI-Africa (2002) *Strategic and Responsive Evaluation of Peacebuilidng: Toward s 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, Horn of Africa Series.

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 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/New Delhi: 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. et al. (1999) *Evaluation: A Systematic Approach*, Thousand Oaks: 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, pp. 285-298.

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, D.C.: National Academy Press, 294-337.

Schirch, L. (2005) *Ritual and Symbol in Peacebuilding*, Bloomfield, CT: Kumarian.

Smith, D. (2003), *Towards a Strategic Framework for Peacebuilding: The Synthesis Report of the Joint Utstein Study on Peacebuilding*, Oslo: PRIO.

Sorensen, N. et al. (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.

Zartman, I. W. (2000) ‘The timing of peace initiatives: hurting stalemates and ripe moments’, *The Global Review of Ethnopolitics*, 1 (1), pp. 8-18.

Table 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 ex-ante	Formative Participatory	No	Insiders with the help of an outsider	Micro	Not specified
Reflecting on peace practice (Anderson and Olson 2002)	Identify lessons learned	Ex-post	Summative Formative	Some	Outsider	Macro	All kinds of initiatives
Paffenholz evaluation framework (in Paffenholz and Reychler 2007)	Review and judge current status to improve and walks user through process and application	Ex-ante, mid-term, ex-post	Summative, formative, participa tory	Yes	Insider/outsider	All levels	All kinds of initiatives
Environmental conflict resolution evaluation (d'Estree and Colby 2000)	Assessing and documenting what has been achieved; Developing criteria and indicators for evaluation	All stages (baseline, mid-term, and ex-post)	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)	Assessing and documenting what has been achieved, identify lessons learned	Ex-post	Summative	Yes	Outsider	Micro	Peace education/co-existence initiatives
D'Estree et al evaluation framework D'Estree et. al. 2001)	Developing criteria and indicators for evaluation	All stages (baseline, mid-term, and ex-post)	Formative and summative	Yes	Insider or outsider	Micro, meso, and macro	Interactive conflict resolution

[1] See Dennis Sandole's chapter 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 Arab-Israeli conflict; Kelman 2005 on the Israeli-Palestinian conflict; Paffenholz 2003, 2005, 2006 on Life and Peace Institute's work in Somalia; Voorhees 2002 on the Dartmouth process).