Chapter 3

Challenges in Peacebuilding Evaluation: Voices from the Field

Mohammed Abu-Nimer

Due to the nature of the implementing organizations, the capacity of the programmers, and how peace work has been traditionally perceived (and many other factors), many peacebuilding programs do not have rigorous monitoring and evaluation (M&E) mechanisms. It is only recently that the field began focusing on the need to be more systematic; such focus has been triggered by both donors and policymakers who began investing more resources in the peacebuilding programs, especially in post-conflict areas. However, efforts in systematizing and evaluating peacebuilding programs have been subject to numerous challenges. This chapter examines some of the core obstacles for evaluators to be aware of in developing and conducting evaluations of peacebuilding programs and projects. The challenges in this chapter were described in detail by thirty-two international evaluators who were interviewed for a wider study conducted by the Salam Institute (see www.salam.org) between 2008 and 2012 (Abu-

Nimer 2016). The chapter begins by examining some of the fundamental obstacles faced by evaluators in developing an evaluation, including those challenges that lie with the donors or peacebuilding programmers themselves, and then moves to more structural challenges faced in conducting evaluations. While the list of challenges discussed is far from comprehensive, the chapter aims to bring to light and emphasize the importance and need for further research on peacebuilding evaluation.

Many peacebuilders in general often see peace work as morally superior to other forms of interaction or intervention in a conflict. In fact, often most outsiders to the conflict (also some people from within the conflict) express a sense of admiration and recognition of the need to work for peace. However, working from a moral advantage cannot compensate for the fact that the majority of peacebuilding implementers stumble and fall short in their attempt to produce evidence to convince the average person, donors, or policymakers that any given intervention has the potential to change the conflict dynamics in a country. As Cheyanne Scharbatke-Church lamented in the early 1990s when meeting with policymakers in Northern Ireland to convince them to shift more of their resources to peace work: "Where is the stuff that I can pound on the table and say, and here's why we know this stuff works better than more guns on the street" (Church and Shouldice 2003). Instead, peace workers have often characterized their efforts as "planting seeds" for future peace, often noting that they do not necessarily expect to view or see the full results of these seeds in their lifetime.

The pressure to produce tangible results has been set by the rapidly advancing technologies, which have dramatically increased the ability of peacebuilding organizations and donors to collect and share data, in turn generating greater demand to demonstrate the impact and the effectiveness of their projects with quantifiable data. In addition, the greater awareness of civil society work and its potential capacity to advance social cohesion increases the demand for peacebuilding programs to produce concrete results (Stern et al. 2012). This is particularly visible in the emphasis placed on empirical evidence by both donors and policymakers; donors are under a great deal of pressure by their policymakers to justify their foreign

assistance budgets, and the trend toward big data has increased their desire for empirical evidence to support their requested levels of funding, prioritizing support for tangible results over peacebuilding needs. Yet a very needs-oriented approach dominates the development field. "We know we all need peace and we easily want freedom, but that's not a tangible."[1]

The need to develop more systematic and widely used M&E of peacebuilding programs is evident. However, the path toward creating strong evaluation systems faces a number of challenges ranging from difficulty in defining peacebuilding programs, lack of evaluator experience on evaluating peacebuilding projects and programs, and donor and programmer behavior, to the very real obstacles faced by evaluators in conflict zones. Often these challenges are interlinked.

When evaluating peacebuilding programs, one must examine a wider scale of influences on the success and impact of peacebuilding programs. Evaluators of peacebuilding programs have to explore linkages with other sectors, rather than just focus on their small and confined efforts of reducing direct forms of violence and advancing peacebuilding. The lack of clear definitions of both peace and peacebuilding, on the one hand, can make it difficult for peacebuilding programmers to create clear indicators of success and, on the other, for evaluators to structure evaluations. Pressure from donors to show quick tangible results clashes considerably with the aforementioned attribution of peace work as toward a moral higher cause and can heighten peacebuilders' resistance to evaluations.

A further dilemma is that peacebuilding work is often subject to two-year program cycles that are dependent on donor funding. These issues are not made any easier by the fact that the majority of the peacebuilding programs are implemented by local partners but designed by international organizations (United Nations, European Union, African Union, Organization for Security and Co-operation in Europe, etc.) or donors (World Bank, USAID, Department for International Development [DFID], etc.), a feature that adds to the complexity of carrying out an evaluation that is relevant to the local partners (Brown 2015). This might also be true of most international development programs, and thus it might not just be a

challenge specific to peacebuilding but international programming writ large. Nevertheless, the evaluator must take into consideration the meaning and impact of such factors in design and implementation on the hope for peace and commitment of local communities to peace when presented by such outside organizations.

Even when evaluators are able to navigate these obstacles and develop an evaluation methodology and tools for a particular peacebuilding program or project, the challenges do not stop there. Often evaluators themselves lack expertise in peacebuilding which can pose an issue not only in evaluation development but also in conducting the evaluation in the field.

Furthermore, conflict zone realities produce certain conditions on the ground which affect people's behavior and responses to everything around them. The proximity of the conflict and its consequences for communities and individuals produce an environment that affects any evaluation design and creates challenges and obstacles for evaluation teams. Identifying and devising strategies on how to handle these possible obstacles is crucial to the success of an evaluation.

CHALLENGES IN DEVELOPING PEACEBUILDING EVALUATIONS

Lack of Consistent Definitions

Definitions of peace (ceasefire, disarmament, new government elections, return to communities, etc.) and peacebuilding (government sector reform, security sector reform, establishment of clear state boundaries, transitional justice, reconciliation, dialogue and education, etc.) range from narrow to broad. Debates and varied usages of these terms exist in both academic and policy circles. This lack of clear definitions creates a challenge when developing evaluations. It, therefore, becomes important to clearly define these terms when developing an evaluation. In this chapter the term peacebuilding is used as an umbrella term that covers all types of intervention that aims to repair or respond to conflict issues in a peaceful

method, such as conflict resolution, post-conflict stabilization programs, peace education, nonviolent direct action, and diplomacy and negotiation.

When evaluating peacebuilding programs, one must examine a wider scale of influence than the immediate effect on individuals, institutions, or communities. Evaluators of peacebuilding programs have to explore linkages with other sectors. The need for broader perspective stems from the fact that peacebuilding occurs in volatile areas with complex situations with numerous factors that can affect the level of impact that any given peacebuilding program can have. For example, without taking into consideration an external factor that leads to a further outbreak in violence, an evaluation could falsely interpret the effectiveness of the program design or the reasons for delays in implementation. This has prompted some evaluators to work to expand the definitions of the field of peacebuilding and of peacebuilders. Michael Lund, for example, argues that it is a mistake to assume that operating in any level of Lederach's pyramid of intervention —bottom-up, top-down, or middle out—makes someone a peacebuilder and thus requires that this person has special criteria to be able to execute his or her specific part of the operation; conversely, being far from a narrow field, peacebuilding often requires the work of individuals in the process who are not peacebuilders (Lederach 1997). Projects targeting civil society groups, security organizations, and military services for providing secure environments in a conflict also have peacebuilding impact and ought to be included in a definition of peacebuilding interventions. The lack of an agreed upon definition of what it means to work for peace or peacebuilding is a primary challenge in both designing and implementing as well as evaluating peacebuilding intervention.

The attempts at opening a wider definition presents further difficulties as to whether or not the project in question is solely a peacebuilding program or if peacebuilding programming is integrated across all of the organization's projects. For example, according to World Vision or Catholic Relief Services, peacebuilding input and design is expected to be a component of all their development projects, in addition to their direct peacebuilding programming. Evaluating these types of programs requires

different strategies. Furthermore, when peacebuilding is integrated into development projects, it tends to be carried out by programmers who are not necessarily trained in the peacebuilding frameworks and therefore have limited awareness of definition, meaning, and strategies of peace.

In addition to the difficulty in defining what constitutes a peacebuilding program, the lack of a clear definition of "peace" in most peacebuilding programs is a major challenge for evaluators. Of the evaluators interviewed for the study, 30 percent mentioned that, conceptually, there is an inconsistent definition of "what builds peace" or conceptualization of what a good peacebuilding program would look like in the field. Without a precise definition of how a particular intervention relates to accomplishing peace on a macro level, it is difficult for evaluators to link the program's outcomes with the way that people, donors, and organizers in the conflict context understand peace. Thus, evaluators can be torn between various conceptual approaches of "what constitutes peace?" As Mary Anderson states:

> We don't really know what constitutes a good peace program, we don't know what makes peace happen in any definitive way. Most assume, if one does peace, then you bring peace. We don't have clear set of benchmarks and that makes it really hard. It's an imperfect field, and it's hard to know. You have to make a good argument for this particular program at this time as having some significant positive impact, or significant negative impact, or no significant impact at all in relation to an unknown state of being, which is the issue then of peace.[2]

The aforementioned statements reflect the need to further clarify the links between how donors define peace work and its outcomes versus how practitioners and policymakers view the meaning and functions of peace work. The further the gap between these three stakeholders in defining "what constitutes peace?" the more challenging it becomes to both measure and capture the impact of peace intervention programs.

Program/Project Design Lacking Clear and Specific Objectives: "Good Things Happened" Syndrome

Another major challenge for peacebuilding evaluation can occur when the program itself is not designed or planned with any clear or systematic view of change. In this study, 40 percent of evaluators interviewed indicated that the programs they evaluated lacked well-thought-out, participatory program design, goals, and indicators based upon conflict resolution and peacebuilding assumptions and principles. Because of financial or contractual constraints and/or lack of professional knowledge and experience, peacebuilders do not always articulate their theory of change and rationale behind their program activities. Another issue is often that the urgent and critical need to respond to the conflict is the primary driving force behind peace worker's intervention. Thus, in many cases, very little effort is invested in a project's design. As stated by an international evaluator (who worked for international donors such as USAID and DFID), in fact, "a lot of cases the program design process was very sloppy, goals were not clear, and no analysis was done.[3] Team members don't have the same shared theory of change, no clear change was ever articulated for the program, and things have evolved."[4]

Many peacebuilding programs tend to lack specific and concrete measures of progress and impact and instead focus on the lofty aim of achieving peace in places. A scholar-practitioner notes, "The problem was not that peacebuilding cannot be evaluated. The problem was that we were not designing peacebuilding programs that were responsive to theories of change."[5] This is only compounded when the evaluator is expected to deliver the judgment or the measurement of programs that operated for years without much strategy of peacebuilding.

Practitioners and evaluators such as Jay Rothman, Marc Ross, and others have argued that practitioners, prior to the start of implementation, need to develop clear and specific definitions of the program's goals as well as a plan for how to adjust such goals due to shifting conflict dynamics. This strategy would

1. help donors avoid imposing externally developed evaluation criteria and goals on programmers and beneficiaries;

2. allow programmers to have ownership of the goals and reduce the level of alienation and resistance among the staff and beneficiaries when dealing with evaluation; and
3. assist in capturing the changes and monitoring the progress in a more effective way through Action Evaluation.

In most programs, the evaluation is usually scheduled solely at the end of the project or with one additional midterm internal evaluation. Limiting the evaluation to the middle and/or to the end of the project rather than already in the development of the program/project design constitutes a major challenge for evaluators in carrying out a useful process that will directly benefit the programmers or the organization. When evaluators only participate at the very end of a peacebuilding project, they miss the opportunity to proactively help the project in clarifying their goals and be more thoughtful and conscious about what they plan (Church and Rogers 2006; Van Brabant 2010). As stated by d'Estrée: "This 'formative evaluation' process and mechanisms can be inserted explicitly in the design phase which will help program managers better plan the design towards achieving their goals. The second objective of being involved early is to simply be able to collect data and form some baseline for comparison, instead of speculating or using proxy measures at the end of the project to capture how things were in the beginning of the project."[6]

Capturing the Impact of Peacebuilding Programs on Macro Peace

How does an encounter program for Israeli Palestinian youth, which has been taking place since 1993, affect the nature of the Israeli-Palestinian conflict and success or failure of peace negotiations? The above is a classic question that is often posed by the average person in the street, policymaker, donor, and even some peace workers. Seeking to link any peace program with the macro-peace processes is one of the most

frustrating and challenging tasks that face peace practitioners. Responding to this question requires many considerations.

Capturing the impact of the peacebuilding programs on macro-peace processes is complicated and challenging. As an emerging field, its practitioners are still stuck focusing on the outcome, results, and monitoring. The peacebuilding approaches, processes, and outcomes do not fit neatly into the logical framework approach in M&E, whose use is influenced by development and other fields. The high level of complexity in peacebuilding programs prevents practitioners from being able to measure tangible outcomes. Peacebuilding evaluators have to consider broadening the range of their impact evaluation in order to be able to maybe capture the potential contribution of peacebuilding programs on the macro peace level.[7] Seeking tangible and causal linkages between peace intervention and macro-conflict dynamics is too narrow a lens to identify the contribution of many peace programs. Not broadening the range of the impact evaluation has a further consequence in addition to the prospect of mismeasurement; evaluations may simply be unable to identify and measure impact. The inability of projects to identify their impact is

> a problem because we are funding peacebuilding programs, and we cannot measure them, or at least, we have a very difficult time measuring them. So, I am worried that peacebuilding is going to start getting a very bad reputation for being too amorphous, too grey, and too fuzzy around the edges and that does not have rigor because it is so multidisciplinary. They are not going to be able to come up with the rigor [design] to actually measure impact that is unique to that multidisciplinary approach.[8]

Identifying Success

A further difficulty in peacebuilding evaluation is how to identify success. While it is logical that donors or international organizations desire macro-level change, successes (and failures) occur at different levels and require different time frames. When describing the differences between measuring

the successes of development programs and peacebuilding programs, Abi Riak, an international development and peacebuilding expert, stated:

> Success [in peacebuilding] is not just [the] vision of people reconciling or hugging in the streets. It's much subtler. Part of the [goal of] evaluation is to understand how to define success and to learn how to talk about success so people understand what success is in peacebuilding. Maybe some people reconciled [because of the project], but macro changes are not taking place.[9]

While it does not mean that macro-level change goals and measurements should be set aside, it is valid to question whether it is realistic to expect macro changes in beliefs as a result of an individual peacebuilding programs. In addition to the constantly shifting context, there is a delayed effect with peacebuilding interventions, especially for training workshop programs. The example that is often suggested by peace practitioners in the Israeli-Palestinian context is the difficulty of relying on attitudinal changes that are measured immediately after three days encounter versus measures conducted after six months, five years, or even twenty years later.

Attitudes, beliefs, behaviors, convictions, and societal structures are extremely difficult to measure and the evaluation tools resorted to for measuring behavioral change are often inadequate for truly demonstrating macro-societal impact. In this study, 30 percent of the evaluators interviewed mentioned that the peacebuilding field lacks relevant tools to measure the aforementioned human processes and structures. Furthermore, the change in attitudes and belief is hard to maintain and sustain over a longer period. It is also hard to change the ideological foundation of a person. A person may walk out of a workshop, dialogue intervention, or other peacebuilding intervention with a change of conflict outlook; however,

> if beliefs can be changed in a five-day workshop, they can be as easily changed back. They can be irradiated, done away with, by the contextual forces. Look my grandson returns from a dialogue. He comes to a home, which is a peace-loving home, but then he goes to the classroom in school, and he encounters an environment where he is called a "traitor." The same is true of his peers, the Palestinian peers, who represent their classroom, or our own neighborhoods and

face an entirely different social context which negates what he has acquired in the wonderful peacebuilding workshop. In terms of collective narratives, he (the participant) brings for them (his peers) an entirely new fresh baby narrative. But, when he returns to his natural environment, with the media, politicians, etc., he faces a bold, very strong veteran collective narrative, and his baby narrative, the one he just required during peace work, stands a very slim chance of surviving in the face of the old, consensually collective narrative.[10]

If we accept such a premise, what should our realistic expectations be of a "one time" peacebuilding meeting or even a two- to three-year-long project? As Gavriel Solomon argues:

The Catholic IRA terrorist from Northern Ireland who, after 17 years in jail, becomes a peace promoter is a rare case. Usually participants will not come away with such extremely new and radically different views. So, how does this compare with a weekend of a meetings between Jews and Palestinians? What chance does a program have to change deeply held convictions? So, whoever believes that total transformation is possible, I don't know where he lives, certainly not in our region, or Northern Ireland, or Kosovo.[11]

Methods

In responding to this challenge, we must not rely exclusively on quantitative data collection or analysis; such data often fail to capture the nuances of peacebuilding work, especially on sensitive change issues (e.g., national or religious identity framing changes).

Relying exclusively on surveys will often not reveal the full story of local people's perceptions of a peacebuilding project. Furthermore, surveys might actually offend or contradict local people's viewpoints on what needed to be done and how it should have been done. For example, in an evaluation of a Burundi conflict resolution workshop, participants refused to fill out the neatly designed survey. Instead, they insisted on face-to-face meetings with the evaluators to narrate their stories.[12]

Thus, as indicated by many scholars and researchers on the need to use mixed methodology when searching for empirical evidence, the case is the same in peacebuilding evaluation (Bamberger 2012). The need to combine

quantitative and qualitative measures in evaluation is overwhelmingly supported by the peacebuilding evaluators. "One very important recommendation for evaluation is you should not be satisfied with quantitative findings; you have to go with qualitative interviews and observations, too. If your data is from the previous morning, do not open a bottle of champagne. Wait for two months, repeat the measurements, and if then you find that your positive outcomes are still there, then write your mother a good letter."[13]

Too Narrow a Focus

Another challenge facing evaluators trying to connect peacebuilding programs with macro peace is that majority of these programs have the tendency to focus on local cross-community dialogue and conflict resolution programs, without looking at the wider environment of the conflict. Organizations in conflict zones need to rethink their strategies of peacebuilding by examining the macro level of analysis. Local-level impact may be occurring due to peacebuilding initiatives, but the situation may be the reverse on the macro level. The reality is that "no local level impact is going to be long lasting if you're going to be in a broader environment that's deteriorating," argues Robert Ricigliano, describing Mercy Corps' work in Afghanistan and their attempts to conduct a wider conflict assessment and analysis to demonstrate the impact of their projects. "The guiding questions were: how can we get a greater macro impact for our programs? What do we need to have in place to achieve such an outcome?"[14]

Raising such questions are crucial in the field of peacebuilding today, especially among the main peacebuilding donors or funders, such as the U.S. Institute of Peace (USIP), USAID's Office of Conflict Management and Mitigation, the Canadian International Development Agency, and the United Nations Development Programme (UNDP). In the case of USIP, for example, it is important to ask, to what extent have its programs in Israel-

Palestine contributed to the macro peace between Israelis and Palestinians? What type of contribution has the organization had in advancing macro peace and stability in Iraq and Afghanistan? To what extent has it had an overall capacity to examine and conduct assessments of macro policies in these regions to be able to evaluate the effects of the organization's projects and programs?

THE PUSH AND PULL FACTORS FOR AND AGAINST EVALUATION DEVELOPMENT

"The Donor Made Us Do That"

More and more donors are demanding more than a list of outputs and input activities and participants. The rising trend is to seek indicators of success and effectiveness and to identify the wider impact of peacebuilding programs. While many peacebuilding programmers may not be interested in conducting an evaluation, they are forced by push factors such as the requirement of donors to provide empirical evidence or to produce a report showing progress, impact, or success as a condition for continuing or sustaining funding. Pull factors include a greater access to more donors, grants, and longer funding to be able to complete the project or program, and practitioners pulled by the desire or need to truly be able to capture evidence showing they are effective beyond the small group of participants. Thus, the general attitude of the staff is often that evaluation is a burden or requirement from the donor, rather than a tool for learning or improvement of the design and implementation. Some staff members do not view the evaluation as instrumental to improve their intervention; on the contrary, they invest some efforts in painting an exaggerated positive picture of the project, by manipulating the evaluators, target audience, or evidence. This attitude often increases pressure on evaluators to focus on "self-serving and glowing reporting." For example, in several evaluations of projects in Sri Lanka, Egypt, and Israel-Palestine, the sponsoring organization requested

two reports: an external document to share with donors and public and the internal document to share with their staff and board members.[15]

Such behavior is influenced by the structural factors of limited resources and the need to pursue funds, especially in the form of governmental contracts. Thus, there is a chronic problem of needing to get money from donors, having adequate evaluations, and needing to ensure the next grant. Under these conditions, it can be difficult for peacebuilding organizations to see evaluations as a learning tool rather than a marketing tool. Similarly, donors are not exempt from a failure to recognize the true value and role of evaluation in peacebuilding intervention. Although the donors in the field of peacebuilding are trying to be more evaluative, they are trapped in their inability "to come up with all sorts of seemingly helpful measures."[16] Donors also tend to treat the evaluation report as a "box to check" on the form of requirement for making their awards, and they are often not involved in the monitoring or the evaluation process.[17]

Some evaluators argue that donors have a responsibility to be more flexible in measuring the output and impact of the projects in deep-rooted conflict areas in which people have been through great deal of suffering and loss and lack basic infrastructure. According to several interviewees, in the past, before donors increased their interest in evaluation, they used to provide programmers with some flexibility to assume that certain positive outcomes would occur as a result of the mere existence of a peaceful intervention in conflict area. The guiding principle was just to make sure that funds were well-spent, which resulted in giving programmers a significant amount of leeway. However, more recently, donors have become "very specific about what the deliverables were going to be and what the time lines were going to be."[18] Richard Blue argues that this type of donor flexibility is needed when evaluating locally run programs: "Putting organizations through the competitive open bidding process is mindless when you are talking about the kinds of folks who have been through 20 years of hardship, as in Liberia, and somehow expecting them to come up and start creating social impact. They can't do it and they shouldn't have to

do it."[19] In peacebuilding evaluation design, imposing external criteria, frameworks, and indicators can be an absolutely disempowering process.

The Curse of the Two-Year Project Cycle

Projects in conflict areas tend neither to abide by the donor's expectations nor by those of the implementing agency. The constantly changing dynamics of the operating environment causes a project's progress to move forward, backward, and sideways, complicating the task of predicting a clear hypothesis for input and change. This fluidity subsequently complicates the evaluation process. For example, in its first year, the Christian-Muslim interfaith dialogue program in Egypt, carried out by the Christian Evangelical Social Services in 2010, was marked by great success and it had a promising potential in reaching out to policymakers and religious leaders. However, in the second year its success criteria completely shifted when the Egyptian revolution erupted and brought down the Mubarak regime. The program was greatly affected due to security, public opinion and atmosphere, and shifts in national priorities.

An intervention peace plan or project is not an instance or a singular event in the conflict context. Peacebuilding occurs in a violent conflict or post-conflict dynamics; thus, there is a need to examine a longer time frame if a project or program is expected to have forward-moving progress and lasting impact, which is often the aim or desire of the donating organization. However, a majority of donors request and expect clear evidence for long-term impact from short-term and immediate designs. As stated by Kinghorn:

> To come in with a project that operates with a project framework that wants a forward-moving predictable, based-on-hypothesis of results, project is just ill-suited and so the two don't usually go well together. Projects don't progress on the year time frames, which is usually what donors want to give or usually the project duration is two years. Yes, usually it's a two-year horizon and so pretty much you need things to be moving around mostly a monthly, quarterly, biannual timeframe. You need to be seeing progress and substantial progress, two to five years depending on the project and because of the backwards-forward nature of

this instable period, it doesn't usually happen. Now if a project was actually moving forward and not encountering any of these conflicts or situations, it could be that it's not—it could be that it's just not doing anything, it's not effectual, it's not actually addressing this instability or the conflict that it needs to address… . That it's not actually getting in there and addressing the system that needs to change, it's marginalized. And so because programs are so locked into this project delivery paradigm that's been set up, [the staff] is usually frustrated because they are not getting their results.[20]

The limited project cycle approach is often based on a result-based approach, which poses the challenge of focusing on the so-called end results, often without taking into account more subtle indicators that could show impact. One of the evaluators interviewed advocates for a systemic approach rather than a result-based approach:

Due to the complexity and fluidity of the conflict situation, it is not possible to follow results only and determine success and failure based on these results. Evaluators ought to consider the whole system of a conflict and to what extent they are engaging in the system of conflict through the peacebuilding project. One can fail in achieving certain results but still engage the conflict system. A country director can get calls from the ministry officer and be blamed for certain outcomes of lack of results, however these can be indicators that the project is engaging the conflict system and leaving certain marks on it.[21]

Lack of Donor Interest in Evaluations

The assumption that "good things will happen as a result of peacebuilding work" is held not only by many practitioners but also by certain donors. As perceived by a number of evaluators:

Donors who are careful about evaluation want to know how it was spent and look for clear indication that "good things happened." Nevertheless, it is not obvious that the donors are well-versed in what a good evaluation method is itself and how they should judge whether what we are hearing, for example, whether a survey done at the end of the workshop is really the right [form of] measure.[22]

It is also more challenging to conduct any systematic evaluation with accurate indicators of success when neither the practitioners nor the donors themselves are specific and clear about their expected outputs and impact.

In some cases, donors have added to the challenge of developing evaluations by not only not developing more specific goals for measurement of success but also having the lack of interest in doing so. For example, in the process of reconstructing New Orleans, a foreign donor gave a large sum of funds to support the sociocultural and structural development of the city. However, when evaluators approached the donor to examine the impact and outcomes from their funds, the donor said clearly and bluntly: "I do not care what they have done with the money. I gave it and that is my duty" (informal meeting with a donor representative in Washington, DC, at American University 2008). The donor was referring to his faith commitment as a primary motivation for why he gave such a large amount for humanitarian relief aid in such context.

Some other possible explanations for the lack of interest in peacebuilding evaluation among certain donors include the following: the ideological or political commitment to support this type of work; the program is too small for the donor's operation; the donor might have a low bar or expectations from the implementers or the design; or a lack of awareness of the need or importance for evaluation. The general perception among such donors and even some practitioners that it is extremely difficult to evaluate peacebuilding programs, especially dialogue programs, can hinder a donor's willingness to fund valuable programs due to the lack of empirical evidence. The difficulty is tied into several of the previously mentioned challenges such as the short program cycle, measuring behavioral change, and the wide range of factors that can affect peacebuilding processes. Some donors may then react by shying away from donating money to peacebuilding programs, which only exacerbates the problem of short program cycles and difficulty in achieving macro-level impact. One international evaluator elaborates on this frustration:

There is so much noise about the outcome variable, there is so much noise about the measure, yet what you have is meaningless. There is no way to know if there is an effect. Does that translate into "you should not put any money in this?" No, it actually does not … then what you probably want to evaluate is the design of the workshop itself… . If you feel confident enough that you have done good stuff to start out, even if you cannot measure, if it is impossible to measure, you may still continue to do it because you have confidence in the setup … [and]what you can do is to be very confident that you have put the right seeds in place and then in theory it should work.[23]

Resistance to Evaluation/Evaluators

Program staff may also be resistant toward evaluation due to the general perception that outsiders are coming just to look for limitations, deficiencies, and ineffective factors. Thus, evaluation/evaluators can have an "oppositional framework" even before it starts.

Programmers may assume that evaluators are hired by the headquarters (HQ) office with the main goal of identifying deficiencies and shortcomings of staff. Thus, the tension between home office staff and field workers may be heightened during the evaluation process if the evaluation and HQ teams do not properly address these assumptions. This dynamic is dysfunctional for peacebuilders whose projects are based on partnership, trust, and relationship building. The typical evaluation approach of an outsider team spending a few weeks examining files and data contradicts the participatory (bottom-up) typical peacebuilding approach. "It's a fear that the people running the evaluation are looking for something wrong with the program as they come in. They talk to people and distribute these instruments."[24]

Programmers and implementers tend to resist the notion of evaluation for various reasons and assumptions too. One evaluator noted:

In the ideology-driven peacebuilding field, many practitioners are doing what they do because they believe in it and believe that they are doing something good and necessary. So, they resist the evaluators who might question the effectiveness and the need or necessity of such a program or intervention. The notion of doing evaluation seems like a waste of time and resources, where practitioners question

why they need to do [an evaluation] when they already know what they are doing.[25]

While lack of clarity of goals is problematic for an evaluator, and the evaluators may find themselves retrofitting goals and as the program evolves readjusting them, however, when program staff is resistant even this can be difficult. Evaluators then have to engage the program staff in a discussion to determine what they are trying to achieve and how they think that they were going to achieve those goals, and that conversation initiates a discussion among program staff about their theory of change. This formative evaluation becomes an actual intervention in the program and, if done well, it can help implementers clarify problematic design issues and assist in adjusting to the constantly changing context and may lead to become more effective over the next period. However, for some peacebuilding officers/managers, this methodology represents a challenge because the evaluator is perceived as an outside consultant who is acting as an intervener in their program. Experienced evaluators have developed the capacity and skills to assume the two roles simultaneously: a program development consultant and an evaluator. Obviously, this evaluation role or approach does not correspond with the principle of "objective evaluation" where there is a clear and strict separation between evaluators and program development consultants.

Evaluators in this study suggested various strategies to overcome resistance to evaluations, and those included:

1. *Pros and cons approach*: Offer the programmers an appropriate and tailored description of the pros and cons of the possible evaluation approaches that they could utilize to evaluate their program as opposed to imposing a set of requirements for one type of evaluation that was determined by either the donor, HQ, or the evaluation team.
2. *Invite staff and beneficiaries to the evaluation team*: Involve some of the staff and beneficiaries in the actual evaluation process (not only selecting the design or approach), especially in design, data collection, interpretation of results, and recommendations.

3. *Adjust your language*: Avoid the use of evaluation jargon to reduce potential anxiety for many programmers who are not familiar with the process and might be overwhelmed by its technical aspects.
4. *Acknowledge past mistakes*: Burnout from previous external evaluation reports is a common source of resistance for a new evaluation. Acknowledge the range of possible mistakes in past evaluations and share previous reports to help reduce possible anxiety or frustration.
5. *Be sensitive to local people's fears and worries*: An external evaluator will be able to gain more entry if they acknowledge local people's worries or fears about evaluations. For example, in an evaluation of a project in Mindanao, one evaluator shared evaluation reports that a lead evaluator in another Muslim community produced with the Muslim beneficiaries of the project. In response, the leaders of the community asked directly: "Did you do work in other Muslim minority contexts? You know we Muslims who live among Christians are always threatened by outsiders who can come to our community disguised in many hats" (Interview, leader of evaluation team in Mindanao 2013). Their response demonstrates the fear of Christian missionary groups among Muslims in South Asia.

CHALLENGES IN THE FIELD—CONDUCTING EVALUATIONS IN CONFLICT AREAS

Scarcity Concerns

Conflict produces scarcity of resources and a strong sense of victimhood among people. In addition, there are often a plethora of donors and development agencies who are constantly struggling (and competing) to meet the needs of the affected populations (OECD-DAC 1991). Thus, the arrival of an evaluation team or a single foreign evaluator into this environment may raise expectations and cause community members and individuals to ask, what are we going to get out of this? To get accurate information and responses, an evaluator must convince project beneficiaries

or wider stakeholders that he or she has no additional immediate funds or aid. Additionally, to elicit genuine responses, evaluators must reassure local people that any negative views that they express about the project will not necessarily result in ending the aid that they are receiving from the agency or organization.

Shifting Realities and Context

To evaluate the impact of peacebuilding program, it is crucial to consider the context in which the program has been implemented. Capturing the context through an early conflict assessment is certainly a helpful mechanism or necessary step; however, in many conflict areas the context is constantly shifting and the conflict dynamics and intervention priorities often change. Accounting for these shifting dynamics throughout the three years or five years of the project's life cycle is certainly a challenge. In this study, 33 percent of the evaluators identified that the complexity and dynamic of conflict environment make the attribution of any macro or even micro impact more difficult.

Except for Action Evaluation methodology, there are few well-defined tools or measurements that can systematically capture the effects of the shifting context on the peacebuilding programs design and implementation (goals, priorities, effect, and overall impact of the project during its various phases and development) (Ross and Rothman 1999). Evaluators attempt to contextualize their findings and report on impact by establishing a timeline or a list of chronological events that took place during the project. Additionally, they look at how certain events affected the project design and outputs. For example, the Office of Transition Initiatives (OTI) at USAID allocated certain funds for supporting Sri Lanka in its transition from war to peace in 2002. Local partners and programmers were encouraged to implement combined design by integrating peacebuilding and development projects. However, when the 2004 tsunami hit Sri Lankan shores, there was devastation particularly in the conflict-affected areas. The OTI budget was raised from \$3–\$4 million to over \$20 million. In 2007, while the

government launched its new war campaign against the Liberation Tigers of Tamil Eelam (LTTE) (after the negotiation collapsed), the final evaluation report had to address questions such as did the OTI operation in Sri Lanka contribute to the peace process? How did the tsunami catastrophe influence the OTI project in Sri Lanka? What effect did such an event have on the objectives of contributing to the peace process?

Because conflicts are not static events but ever-changing organisms, often evaluations can only detail a project's impact in a specific moment of time. Evaluations, especially those designed only for a singular short-term project, highlight the impact right at that moment, but usually are not designed to analyze effects after the end of the project. However, peacebuilding projects in their nature are meant to have longer-term effects that are often either not felt or not fully embedded right at the end of the project.

> You can be satisfied with an evaluation only the morning after you have completed the workshop. There is a delayed post-test effect. After the evaluation is completed, there may be more than one violent attack, and in the months in between, many political issues and tragedies, which could have nothing to do with peacebuilding. You have no control over [these events], and most everything remains the same. However, in regions like [Israel-Palestine] events happen twice a week and what you get is a delayed post-test that may just reflect the recent events, more than anything else.[26]

In addition to the shift in macro events, there are micro changes on the ground that can affect an individual's life and determine the extent of effect of a peacebuilding project on his or her life. Most evaluation designs are unable to capture the effects of these events in any systematic way.[27]

Stressed Program Staff

Operating in a conflict area generates more stress and frustration among programmers than in other contexts. Thus, there are certain challenges related to the fact that in a conflict area

although, [project staff] might not be traumatized in the same way as violent trauma victims, nevertheless, they are completely stressed out, because the project usually doesn't go well [as planned]… . A lot of the staff that I have encountered, as soon as I go into this they wanted to justify their stance, [explaining] what happened, what went wrong and whose fault it was. [staff members], were on a justification standpoint or "we are trying so hard," "they don't appreciate what we are doing," "look what's happening" [managers/external or HQ] are unable to see this.[28]

The lack of recognition of the impact of conflict dynamics on the capacity of peacebuilding projects to demonstrate concrete and tangible impact often leads to frustration among field officers and even leads to certain level of resistance to result-based approach as opposed to a system approach.

In this study, a few evaluators pointed to excessive staff stress due to a combination of a donor-imposed accountability paradigm, the conflict environment, and self-preservation issues as a challenge for evaluation.

[Program staff have a] professional investment in these projects. It's their job, this is how they spend their day. This is the majority of what they give their focus and energies to. So, when the project isn't going well, and they are not getting their results or they are getting conflicts because of their good work, they get completely stressed out and frustrated and angry.[29]

An evaluator's approach ought to take staff attitudes and frustrations into consideration, especially with the data collection methods and design. Mechanical and detached social science approaches tend to produce more alienation among the staff on the ground and can increase the resistance to the evaluation.

ACCESS: SECURITY AND TRUST

Trust Concerns with Government or Security Force Involvement

Lack of security and distrust of others are two major obstacles in any conflict. As one Palestinian participant in Ramallah responded when asked

for his honest opinion of evaluations: "We live in a reality in which you cannot trust anyone in your environment. Everyone might hurt you and threaten your sense of security." In this study, 27 percent of the evaluators identified that security concerns were a significant obstacle to conducting peacebuilding evaluations, and 27 percent of the evaluators specifically mentioned that dishonesty and mistrust can skew data results. Furthermore, 30 percent indicated that access to data, beneficiaries, or simple physical mobility inhibited the success of the evaluation.

Lack of trust in a conflict context is a product of the state of relations among people in conflict areas that influences relations with evaluators as well. Evaluators can offer reassurances that the evaluation is a confidential or anonymous process, but many people in a conflict area may not believe such assurances.

> Thus, evaluators have to gauge how much they can take what people say at face value by [determining] the general level of trust and confidence in the community or the group and [by considering] whoever is being interviewed or surveyed. If there is a high level of mistrust, the chances are there is going to be higher levels of not necessarily dishonesty but not complete honesty or complete information sharing.[30]

The presence of government agencies, especially security forces, undermines the security of evaluation participants. For example, in an Egyptian focus group discussing the impact of an interfaith dialogue program, the participants could not speak freely about the conflict or group relationships due to the daily presence of the security forces in their meetings and focus group discussions. Sitting at the end of the room, a government security officer took notes while evaluators posed their questions about the conflict and Muslim-Christian relationships in Egypt.[31] In response, the participants and the evaluation team had to develop certain codes and signals to communicate about sensitive issues that the government might perceive as political or problematic. The evaluation team then shifted their approach, using private, individual interviews to elicit more genuine and revealing information from the participants. Similarly, in

Sri Lanka, a representative from the LTTE forces sat listening in the training room while evaluators posed their questions to measure the impact of a peacebuilding project in war zone areas in Baticaloa. Cognizant of the LTTE representative's presence, the focus group members began to praise the LTTE's operations while criticizing the Sri Lankan government's policies in the region. However, when the evaluators shifted their methodology and conducted anonymous interviews, participants felt freer to voice their criticism of the LTTE's impact on the project.[32]

Furthermore, government security forces and armed groups can affect participation in evaluations by taking a step further and trying to include their representatives in the evaluation team. "When the security forces in a Central Asian state suggested a facilitator to manage the list of contacts for interviews, this was a threat to evaluators and their interviewees from participating or collecting data."[33] In response, it was crucial to get a local researcher who was not connected to security to be the facilitator.

Conflict Group Dynamics

The asymmetric power relations and the various conflicting and nonconflicting groups' interactions in any conflict situation affect how the individuals and their collectives perceive their relations and what constitutes a successful intervention. Thus, when the evaluator is engaged in the process of capturing success and failure of the intervention, one of the obstacle or challenges is to diagnose, name, and identify the effect of the conflict group dynamics on the responses from each participant.

Managing focus groups in conflict areas can also become sensitive due to the conflict dynamics. Underlying ethnic issues and potential triggers require local facilitators to detect or avoid them in order to elicit genuine responses from participants. Such skills are not a necessary part of the typical "professional/researcher evaluator profile."

Evaluators must be sensitive to the extent to which the data collection methods and evaluation designs are affecting the power dynamics of the conflict. For example, a focus group discussion with Tamil and Sinhalese in

Sri Lanka during an evaluation in 2010 could not be carried out due to the security restrictions and confidentiality concerns. The evaluation team had to find creative ways to gain access to the beneficiaries (phone interviews, a dinner party, an invitation for a wedding reception, etc.). In another case, data regarding the peacebuilding projects involving the Muslim Peace Secretariat in Sri Lanka had to be gathered through more confidential tools of data collection due to the sensitivity of the content as security services were monitoring and blacklisting outspoken opposition members and there was an overall threatening environment. Surveys were also not possible due to security. Evaluators in this case chose to conduct individual interviews as the primary data collection method because they ensured the highest level of confidentiality and trust for the interviewees.[34]

Mobility and Access

Security factors often prevent evaluators from reaching peripheral or less urban areas of a country (which are often less secure). Thus, evaluators in conflict areas find that they need to make sure that the sample of interviewees or stakeholders are more inclusive and pay special attention to the peripheral areas where security conditions are less conducive to travel or even limited to only basic communication for linguistic and/or security reasons. To reach a rural community while conducting an evaluation of a project in northeastern Sri Lanka in 2009, one of the evaluation team members had to drive eighty miles into the nonclear areas (LTTE-controlled area with no government troops or security). This evaluator found that, in fact, in any LTTE-controlled Tamil area, evaluators and programmers could not access the local leaders without explicit permission from LTTE, thus limiting the type of leaders to which evaluators had access.

The limitation on travel often limits the scope of the evaluation to investigate the program's impact on only a certain group of beneficiaries. As Anderson stated: "When we were doing the OECD evaluation, [we were limited to meeting with people who were] mainly in the capital city and then only people that speak English, anyway. You really don't get a

perspective if you are staying in only talking to one side or one group, or one echelon of society."[35]

Mobility into the conflict area is also problematic when the conflict area is large and covers most of the country. Evaluators will not be able to get to certain parts of the country, in particular affected areas, and typically are only able to focus on the most accessible areas. It is very different to work on peacebuilding projects in a small, accessible country than in a large country lacking extensive infrastructure.

> In Macedonia, it was fairly easy because Macedonia is small place and most of the conflicts occurred pretty close together. But in a place like Liberia where a lot of the conflict was in the big patches of the country, we had trouble getting around because of bad roads and everything else is that geographically you are constrained in the same way. In Sri Lanka, we were constrained by the conflict from getting to the places where there was active conflict or where conflict was still prominent.[36]

Local evaluation participants also face security concerns. In Pakistan, a group of fifty project beneficiaries could not meet in their region due to the threat of Taliban forces who opposed their Madrassa reform project, and they instead had to travel for two days by train and bus to Islamabad to discuss their Madrassa teacher training project and its impact on their community.[37]

In a conflict zone, the identity and the presumed religious affiliation of the external evaluators can restrict their access to entering and conducting a peacebuilding intervention in a country as a whole or specific region. There are certain areas in which some external evaluators cannot have access without a local person acting as a facilitator. In Pakistan, while conducting an evaluation of a Madrassa teacher training project, it became impossible for the team of Western external evaluators to access the project's staff and beneficiaries operating in the Pashtun areas. The team then delegated a local religious clergy who had previously worked with the Madrassa project to manage the relationship and gather the necessary data. In this case, a Western, Christian external evaluator would have not been able to gain access to any of these schools or teachers.

Fear and Distrust of the "Outsider"

Security is not the only obstacle that prevents access to data collection or interaction with participants. Suspicion and distrust of outsiders is often a major challenge to be overcome by foreign evaluators.

Access to certain social sectors may be an issue also. For example, in the case of evaluating peace educational programs in three countries (Armenia, Indonesia, and Albania), the evaluators could not meet parents, due to timing and other intervening factors. Thus, the entire evaluation of peace education and its impact on the children were completed without direct input from the parents (Interview with Carolyne Ashton, 2010).

Even without the constraints of security and risks for peacebuilding programmers, many local partners feel a certain degree of alienation from and distrust of external foreign evaluators that requires evaluators to work in building relationships to be able to gain even minimal access to information. Building such relationships is a crucial part of the external evaluator's role in a conflict area. Overcoming the cultural and language barriers is the first task. However, that is not sufficient by itself; connecting with the staff and beneficiaries with respect and dignity is the gateway for getting relevant and reliable data. Unfortunately, in many cases evaluators do not have more than a few hours in the community to collect data (survey, interviews, or focus group discussion), a time that is insufficient to build trust, rapport, or any form of relationships with the beneficiary or local staff. As a result, the responses are kept on the formal and surface level.

Lack of Baseline Indicators

The field of peacebuilding is still struggling to develop its generic set of indicators of success and clear and systematic baselines for measurement of indicators of peace (van Barbant 2010; Abu-Nimer 2003; Lederach, Neufeldt, and Cuthbertson 2007). There is growing awareness of the need for programs operating in conflict areas to take the time and allocate the necessary resources to develop their baseline data prior to their starting

point.[38] In this study, 40 percent of evaluators agreed that there is a lack of appropriate tools, for example, for collecting baseline data and for holding real focus groups as challenged by working in a conflict area.

The difficulty of developing a set of indicators and baseline data for a long-term evaluation also stems from the fact that peacebuilding more is political than many other areas of intervention, for example, health or other branches of development. It also tends to be more convoluted and nonlinear. "I think [peacebuilding projects are] about politics at some level or another even if it does not appear on the surface to be political. So, you are really talking about finding the spaces where you can do things, so if you start planning for year one and you are laying out things for the next five years, it is not going to be very linear."[39] Opportunities emerge or develop throughout the course of a peacebuilding project as a result of the fluid nature of the conflict and its politics. Therefore, when programmers decide to shift direction and change their course of action to capitalize on that window of opportunity, it is then difficult for an evaluator to develop baseline data or follow a set of standard indicators of success. At least 25 percent of the interviewees in the study indicated that the political nature of peacebuilding increases the inability to function due to domestic and international government restrictions and/or negative response to this work, that is, complexity, structural problems, and intervening variables.

Finally, the heightened concern for the security and safety of the staff or programmers might influence the choice of data collection. The higher stakes of succeeding in a conflict context can also require that evaluators pay more attention to how their evaluation will benefit the programmers in the field. An evaluator's decision to include certain groups and exclude other stakeholders from the data collections can influence and shape the results of the evaluation. For example, U.S. government–funded projects and evaluators are prohibited from contacting groups defined as terrorists— for example, the LTTE in Sri Lanka and Hamas in Palestine. Despite the fact that these groups and their constituencies have been affected by the specific peacebuilding projects implemented in their area of control,

evaluators and programmers are not allowed to contact the leadership affiliated within these areas.[40]

CONCLUSION

Obviously, the aforementioned list of challenges is selective and noncomprehensive. The purpose of the discussion is to highlight the need to place such challenges on the current and future research agenda of the field of peacebuilding evaluation. While this chapter explored challenges that face evaluators of peacebuilding, nevertheless, some evaluators who were interviewed for the study argued that the aforementioned set of challenges can also be encountered in evaluating development, education, or/and other area of social and economic development and not only in peacebuilding. Therefore, they consider that there is no need to identify a special set of challenges for peacebuilding. However, the serious risk of supporting this argument among donors, peacebuilding, and development agencies is the deployment of international and national evaluation teams to examine the impact of peacebuilding programs. Such evaluators tend to use cookie cutter templates without fully understanding or being equipped with proper tools or qualifications to understand the impact of the conflict and its dynamics on implementing peacebuilding programs. The results can be devastating to peacebuilding and donor organizations which, in many conflict situations, devoted the career and lives to promote peace and coexistence in their communities or society in general.

This is not a water tank project in a conflict area in which the evaluation team is examining if it was used efficiently or built according to certain technical regulations. For peacebuilders and their beneficiaries or stakeholders, the stakes are very high if the evaluator/s fails to capture the impact of the immediate or long-term effect of their efforts on the promotion of peace. Thus, responding to the aforementioned challenges and taking them into consideration in a given peacebuilding context become

more urgent factors in the design, implementation, or follow-up of any evaluation.

NOTES

1. Interview with Abii Riak on July 6, 2009.

2. Interview with Mary Anderson on September 12, 2009.

3. Such a conclusion was also noted by Blum and Kwano-Chiu 2012.

4. Kupperstein interview on April 5, 2010.

5. Interview with Cheyanne Scharbatke-Church on April 28, 2010.

6. Interview on March 10, 2010.

7. See also studies on the need to broaden the design and methods to capture macro impact: Stern et al. (2012).

8. Interview with Cohen Wood on April 20, 2010.

9. Interview with Abii Riak on July 6, 2009.

10. Interview with Gavriel Solomon on August 20, 2009.

11. Ibid.

12. Interview with Amr Abdallah in 2011.

13. Ibid.

14. Interview with Robert Ricigliano on July 8, 2009.

15. The defensive attitudes against evaluation can also stem from the situation in which the donor is not clear about the objective and possible expectation from the peacebuilding project. A former international development officer stated: "On a particular project that I worked on in Pakistan, the donor was always changing their mind about what they wanted the project to do, and the project started and stopped a lot of initiatives as a result. We also felt really defensive about some project failures that were a direct result of donor policy (for example, the project didn't complete any activities for 3 months while the donor refused to obligate more funds while they were deciding if they were going to cancel all contracts with international organizations and shift all funds to local organizations)" (Interview, Barday 2017; see also Allen et al. 2014).

16. Interview with Robert Ricigliano on July 8, 2009.

17. According to an interviewee (international evaluator—requested to be anonymous) who worked with U.S. agencies in Iraq between 2003 and 2010, an organization in Washington, DC, that awards over $2 million in grants required two pages from each awardee. In addition, they would never contact the fellows for any follow-up once the grant is completed and the reports are filed.

18. Interview with Robert Ricigliano on July 8, 2009.

19. Interview with Richard Blue on October 5, 2010.

20. Interview with Kinghorn on November 16, 2009.

21. Ibid.

22. Interview with Malhotra on January 20, 2010.

23. Ibid.

24. Interview with Ross on September 21, 2009.

25. Interview with Nike Carstarphen on February 20, 2010.

26. Interview with Gavriel Solomon on August 20, 2009.

27. Such impact was also captured in more details on the study of OECD DAC (2012); CDA Collaborative Learning Projects (2004).

28. Interview with Kinghorn on November 16, 2009.

29. Ibid.

30. Interview with Nike Carstarphen on February 20, 2010.

31. Based on experience of the author while working in Egypt between 2000 and 2016.

32. Based on the experience of the author in Sri Lankan context between 1996 and 2010.

33. Interview with Chetwynd on February 10, 2010.

34. Based on the experience of the author in Sri Lankan context between 1996 and 2010.

35. Interview with Mary Anderson on September 12, 2009.

36. Interview with Richard Blue on October 5, 2010.

37. Based on the author's experience in conducting this evaluation in Lahore and Islamabad in 2010.

38. Interview with Diana Chigas on September 10, 2009.

39. Interview with Richard Blue on October 5, 2010.

40. For example, see one of the lists that identifies many organizations that cannot receive any funds from European Union: http://www.consilium.europa.eu/en/templates/content.aspx?id=26698.