

Evaluating Peacebuilding: Not Yet All It Could Be

Cheyenne Scharbatke-Church

1. Introduction	460
2. An Abridged History of Evaluation and Peacebuilding	461
3. Mapping the Nexus Between Evaluation and Peacebuilding	463
3.1 What Gets Evaluated?	463
3.2 Who is Involved in Evaluation?	463
3.3 Norms of Current Peacebuilding Evaluation and Evaluation Best Practice	465
4. Peacebuilding Evaluation: What Makes It Different?	468
5. Open Questions and Challenges	471
5.1 What is the Purpose of Peacebuilding Evaluation?	471
5.2 Is the Peacebuilding Field Learning from Evaluation?	473
5.3 Does Evaluation Feed Accountability in Peacebuilding?	474
5.4 Evaluation as a Means to Different Ends: Bad Practice	476
6. Recommendations: What Can Researchers, Practitioners and Donors Do?	478
7. Conclusion	480
8. References	480

1. Introduction

Whether an engineer, a doctor, a politician or a mother, everyone practices evaluation every day of their life. One evaluates whether the rice has finished cooking, whether the proposed law will address the problem at hand, which diagnostic tests to do in response to the patient's symptoms and whether the bridge design is adequate for extreme weather conditions. Each of these evaluative acts has a purpose in mind, requires information, and assesses that information against the context and against standards (explicitly and implicitly), in order to catalyse an action. Although everyone has this inherent familiarity with evaluation, in a professional setting it often becomes rife with misunderstandings, complexities and challenges.

Therefore, it is useful to start with first principles – what *is* evaluation? Definitions abound within the professional evaluation field, with many of the major evaluation theorists having developed their own variations (Patton 2008; Rossi et al. 1999). Almost all of these have at their core a set of common characteristics: the systematic nature of the process, competent data collection methodology, and assessment or valuing of the findings. Overlaying these characteristics with peacebuilding, one can define evaluation for our purposes as *the use of social science data collection methods (including participatory processes) to investigate the quality and value¹ of programming that addresses the core driving factors and actors of violent conflict or supports the driving factors and actors of peace* (Church 2008).

In this definition, “quality” refers to the calibre of the implementation; including the conflict analysis from which a peacebuilding strategy is derived, the planning as well as the implementation process (a blend of logistics, tactics and peacebuilding technique). “Value”, on the other hand, inquires into the changes associated with the intervention and their significance to the ‘target’ population in terms of stopping violence or building peace. Both quality and value are essential components of programme evaluation.

This chapter explores the state of the art of evaluation in peacebuilding. After reviewing recent developments (*section 2*) and current practice (*section 3*), it proposes that peacebuilding evaluations are generally not delivering accountability and learning in the manner in which they should for two primary reasons. First, the average evaluation is not grounded in the basics of good evaluation practice. Significantly more attention is given to responding to peacebuilding's perceived ‘distinctiveness’ and the challenges this distinctiveness raises than to ensuring that the basics are covered (see *section 4*). The second reason is that the core drivers of evaluation – accountability and learning – are rarely held at the heart of the process (see *section 5*). *Section 6* gives some recommendations for improving evaluation in the peacebuilding field, followed by a short conclusion.

¹ This definition uses the terms “quality” and “value” as they tend to resonate more with the field of peacebuilding. The evaluation literature tends to use “merit” and “worth”.

Methodological challenges, of which there are many, are not covered in this chapter. While the field needs to address these challenges in a thoughtful manner, the issues of quality and the motivations behind evaluation are at the foundation of evaluation practice. No improvements in methodology will fundamentally change the contribution of evaluations if these issues are not adequately addressed.

2.

An Abridged History of Evaluation and Peacebuilding

The late 1990s and early 2000s saw a surge of interest in peacebuilding evaluation, with numerous international meetings, publications and evaluations. ALNAP, INCORE, and the University of Wisconsin were three, of many, who sought to advance thinking on this issue.² The discourse focused predominately on how one would evaluate the outcomes of projects. In 2003 and 2004, two major studies were published that had significant reverberations on the conversation around evaluation of peacebuilding.

First, the publication in 2003 of Mary Anderson and Lara Olson's *Confronting War*, the culmination of many years of collaborative learning within the peacebuilding field through the Reflecting on Peace Practice (RPP) project, challenged the field to assess their contribution to "peace writ large" (PWL).³ PWL refers to the successful end state that peace practitioners are seeking to achieve and RPP posited that all projects should be able to articulate their contribution to this state.⁴ From an evaluative perspective this raised numerous methodological challenges and dilemmas and it quickly became the 'sexy' topic.

The following year saw the publication of the high-profile Joint Utstein Study of Peacebuilding, led by Dan Smith. One of the primary findings from this broad review was that peacebuilding had a strategic deficit. Of the many recommendations given in response to this deficit, the most significant to evaluation was that: "the [...] evaluation community needs to recognise that impact assessment at the project level is not proving to be viable and to shift it to the strategic level" (Royal Norwegian Ministry of Foreign Affairs 2004, 14).

The combination of these two projects, despite their differing conclusions, contributed to a growing tide of interest in assessing contributions to the strategic level or PWL. These discussions indicated some notable changes in the discourse on evaluation in peacebuilding. They showed a more sophisticated understanding of the complex change process outside of the immediate sphere of project influence. They generated new thinking in terms of what PWL is,

² Tanya Spencer's work for ALNAP in 1998 is one of the first dedicated studies on evaluation and peacebuilding. This was followed by Cheyanne Church and Julie Shouldice and their two INCORE publications (2002, 2003). The University of Wisconsin hosted a 2-day international meeting in 2002 looking at measuring effectiveness.

³ There is some debate about whether or not this is a reasonable expectation to have of peacebuilding projects.

⁴ This is sometimes called evaluation of impact at the societal level.

theories of change and strategy, they increased interest in evaluation and to a degree generated new ways of tackling difficult evaluative challenges.⁵

However, this enthusiasm unintentionally reduced the attention given to what had come to be seen as the ‘uninteresting cousin’, namely outcomes-focused evaluation at the project level. This unfortunately happened before much of the work started in this area in the late 1990s had come to fruition, either through clear methodological advancements or through integration into peacebuilding organisations’ regular practice. Throughout the latter half of the decade, outcome evaluations did continue to sputter along, but the mechanisms for pushing the thinking on them significantly reduced as key people turned to consider the other, apparently more pressing challenges.

As the decade drew to a close, the pendulum started to swing back, and today one can see a resurgence of interest in outcomes-focused evaluation related to peacebuilding. The number of agency-based jobs in monitoring and evaluation has much increased, as have the number of outcomes-focused evaluations. Equally important are the recent spat of new projects that have the potential to contribute to this field. CARE International and International Alert initiated a project in early 2010 that will “test the effectiveness of a number of peacebuilding theories of change and pilot innovative but simple approaches to monitoring and evaluation”.⁶ Search for Common Ground, supported by DFID, plans to launch an online learning portal dedicated to the design, monitoring and evaluation (DM&E) of peacebuilding and conflict transformation.⁷

Furthermore, the United States Institute of Peace (USIP) and the Alliance for Peacebuilding have teamed up to offer a “neutral space for a committed group of donors and implementers to discuss and reflect upon the fundamental issues around evaluation within peacebuilding”.⁸ Meanwhile, INCORE and the Canadian International Development Research Centre (IDRC) are working together to “map out and explore the most vexing questions and issues surrounding the evaluation of research in and on violently divided societies” (INCORE/IDRC 2009, 1).

Work on understanding contributions to PWL has also not stopped. For instance, the USAID’s Conflict Management and Mitigation team are leading an effort on developing seven families of theories of change with associated indicators. All of these efforts are likely to contribute to the ongoing evolution of peacebuilding evaluation.

5 Attempts at assessing contributions of peacebuilding to the societal level have been made by Goodhand/Bergne for DFID (2004) and Brusset/Otto for Search for Common Ground (2004) amongst others; however, there are not yet any methodologies for assessing PWL, nor significant literature available to review.

6 Project overview, on file with the author.

7 Announcement email 26 February 2010, on file with the author.

8 Personal email 16 February 2010, on file with the author.

3.

Mapping the Nexus Between Evaluation and Peacebuilding

3.1

What Gets Evaluated?

At first glance, the question of what gets evaluated seems relatively straightforward. One can evaluate projects, programmes, strategies or organisations. In the case of peacebuilding, the majority of evaluations are conducted on projects. However, on second pass the question can become more complicated, as one finds oneself in the middle of a significant debate in the broader field: what, in fact, is peacebuilding?

On the one hand there is the ‘big-tent’ approach of the UN, where peacebuilding is akin to state-building in a post-war context. This includes a wide range of programming, for example militia demobilisation programmes, community literacy programmes and road reconstruction projects that links communities to markets.

Conversely, there are those who use a more targeted definition, as will be done in this chapter, where peacebuilding refers to any work that seeks to address the driving factors and actors of conflict or peace generation. Such work might include dialogues between feuding villages or community councils established to mediate conflicts before they turn violent. At first glance this targeted approach appears to have clearer boundaries, however, it can still be difficult to say where exactly that boundary lies. For instance, do all measures within transitional justice programming fit this more narrow definition?⁹ Regardless of what one perceives as peacebuilding, it has ramifications on the processes, analytical frameworks and standards that are utilised in the evaluation process; as such it is an important question to consider.

3.2

Who is Involved in Evaluation?

Whether one is in the big or little tent, the agencies who will be subject to an evaluation – technically referred to as the *evaluand* – remain generally the same. The most common is the implementing actor, which includes local civil society organisations (CSOs), international non-governmental organisations (INGOs) and United Nations agencies, such as the UN Development Programme. Government donors also commonly initiate evaluations of the peacebuilding programming they support in a region. It is rare for the government of a country where peacebuilding is occurring to undergo evaluation of their work. However, there is a growing

⁹ The OECD-DAC *Guidance on Evaluating Conflict Prevention and Peacebuilding Activities* (OECD-DAC 2008) uses the peacebuilding palette originally published in the Utstein Report, which does include transitional justice (TJ).

movement led by the UN Evaluation Group¹⁰ for country-led monitoring and evaluation,¹¹ which would include evaluation of the government's work as well.

Although these implementing actors are the subject of the evaluation, they are generally not the parties that catalyse the evaluation to take place (i.e. the *commissioner*). Evaluation started to appear on the agenda of the peacebuilding community at the end of the nineties due almost exclusively to pressure from the government donor community. Things have changed very little since: **government donors are the primary commissioners of evaluations today.** Most commonly, this forms a standard component of grant contracts, depending on a donor's specific criteria (e.g. all programmes that exceed a certain finance commitment must have an evaluation).¹²

Though the peacebuilding field commonly looks to the donor community to drive the evaluation agenda, this is a matter of habit rather than necessity. A few organisations have taken charge of their evaluation agenda and have adopted their own criteria for initiating an evaluation: for instance, programmes that represent innovative models or that have been adapted from one context to another must be evaluated.

In addition to the evaluand and the commissioner, *evaluators* are central figures in the evaluation process. There is a large professional evaluation community that grounds their work in the evaluation discipline; however, peacebuilding evaluators are generally not part of this community.¹³ Instead, they are usually consultants or academics who have peacebuilding backgrounds, with teams composed of national and international members.

Similar to overall peacebuilding practice, to date there is no accreditation process or common set of credentials used to signify a 'professional peacebuilding evaluator'.¹⁴ Consultants commonly gain their experience in evaluation through on-the-job learning, while academics transfer their social science research background to the task at hand. In the international development field they can also join the International Development Evaluation Association (IDEAS), whose mission is "to advance and extend the practice of development evaluation by refining methods, strengthening capacity and expanding ownership".¹⁵

In addition to the three roles of evaluand, commissioner and evaluator, there are three further roles commonly found in evaluations. The evaluation *user*, *manager* and *consumer* all have an important function in the evaluation process. *Table 1* summarises the roles, their function and the actors commonly holding each role. It is possible that each role is held by a different actor and equally likely that an actor holds numerous roles. For instance, a donor may commission an evaluation and be an evaluation consumer while the evaluand is also the user.

10 See United Nations Evaluation Group at www.uneval.org/.

11 See UNICEF at www.unicef.org/ceecis/resources_10597.html.

12 A distinction should be kept in mind between the governmental donor community and the philanthropic community, as commonly represented by foundations. This group has been slow to come to the evaluation discussion, with a few exceptions such as the Hewlett Foundation or the Gates Foundation, who is championing extremely rigorous evaluation methodologies.

13 The evaluation discipline has been in existence since the 1950s in North America and now has national professional associations around the world that promote the development of evaluation practice and policy.

14 For evaluators in general, the Canadian Evaluation Society has established an accreditation process in 2010 for members of its national association; see www.evaluationcanada.ca/site.cgi?en:5:6.

15 See www.ideas-int.org/home/index.cfm.

Table 1: Roles in Evaluation

Roles	Function	Actor
Commissioner	Decides that an evaluation needs to occur	Donor, organisation leadership, programme head
User	Utilises the conclusions, lessons or recommendations in decision-making	Donor, organisation leadership, programme head, programme team
Evaluand	Provides documentation and enables access to information sources	The organisation, programme or office that is the subject of the evaluation
Evaluator	Conducts the evaluation	External consultants, evaluand, internal staff
Manager	Provides oversight and management of the evaluation process with an emphasis on quality	A mid-level person who is not part of the evaluand team
Consumer	Reads evaluations to mine for relevant information	Anyone with an interest in gleaning information from an evaluation

There is one group that is striking in their absence in this discussion so far: the beneficiaries or participants of peacebuilding programmes. The current reality is that this constituency is predominately called upon to be sources of information only. True participatory evaluation, where participants play a key role in every stage of the process, is exceedingly rare. In fact, as will be discussed later in the section on accountability (see *section 5.3*), it is not yet common for peacebuilding agencies even to share the results of an evaluation with those who they seek to aid.

3.3

Norms of Current Peacebuilding Evaluation and Evaluation Best Practice

No two evaluations of peacebuilding are exactly alike, with some using exemplary standards and others indicating a clear lack of evaluative expertise. Though there is a large range of practice, one can draw some overarching conclusions that provide insight into current norms and practices. A comparison of current norms in use by the field with best practice standards from the evaluation discipline indicates that peacebuilding evaluations are not consistently aligning to what is seen as central to quality: feasibility, utility, propriety and accuracy (see *Box 1 overleaf*).

Box 1

Evaluation Standards of the American Evaluation Association

Feasibility:

standards intended to ensure that an evaluation will be realistic, prudent, diplomatic and frugal.

Utility:

standards intended to ensure that an evaluation will serve the information needs of intended users.

Propriety:

standards intended to ensure that an evaluation will be conducted legally, ethically and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.

Accuracy:

standards intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the programme being evaluated.

Source: American Evaluation Association (AEA);
see www.eval.org/evaluationdocuments/progeval.html

The average peacebuilding project evaluation lasts between 15-25 working days and flows through a reasonably standard sequence of steps: project document review, data collection, analysis and report writing. This sequence mirrors a classic research project, which signals a confusion that exists amongst many in the peacebuilding field around the difference between evaluation and research. As illustrated by the large number of academics hired to do evaluation, programme evaluation is often thought to be synonymous with traditional social science research.

Though the two overlap, particularly as they utilise predominantly the same data collection methods such as surveys, focus groups or interviews, programme evaluation is distinct from research in a number of ways.¹⁶ They differ in their origin, purpose, process considerations, stakeholder reactions, standards of data collection, forms of analysis and deliverables. Some of these differences are a matter of degree, whilst others are quite significant. The lack of a clear distinction between the two practices has resulted in a common practice in the field today where mini-research studies are being labelled evaluations.

From the perspective of the evaluation discipline, these studies often fall short of the accuracy standard. Common accuracy gaps include a lack of contextual grounding and inadequate description of the evaluation purposes and procedures. Further, the feasibility standard requires that an evaluation is realistic in terms of what it seeks to achieve within the timeframe given. Peacebuilding projects vary significantly in size, but the time allocated for the evaluation remains quite similar, which raises feasibility concerns.

¹⁶ It should be noted that the overlap between programme evaluation and research increases when one steps away from traditional social science to action research or policy-orientated research.

Within the data collection stage, peacebuilding evaluators are generally expected to gather original data from sources on the ground, as desk review evaluations are a rare occurrence. Typical sources would include project participants, partners, donors, government officials and international agencies; with groups outside of this sphere such as journalists, academics or small businessmen far less commonly included. The practice of getting original data does support accuracy; however, at the same time the incomplete range of perspectives being included raises some questions as to the validity of information which falls within the utility standard.

A new trend is to explicitly state the expectation that the evaluation process itself should do no harm, or be conflict sensitive. This aligns very well with the propriety standard, which requires that the welfare of all involved in the process is given due consideration.

With regards to data, one common challenge for evaluators is that limited data exist to build from due to the absence of monitoring. Monitoring is an ongoing management function that seeks to gather key information to inform decisions in real-time. Generally done by programme staff, monitoring tracks the progress towards change (results), reviews changes in the context against the programme model and assures that programme implementation is going as planned (Church/Rogers 2006). Tracking the number of people who attend truth commission sessions in order to gain a sense of the public's direct exposure to the process would be an example of monitoring. It could contribute to decisions about the need for a larger or more targeted information campaign to inform the public, for instance. This data could also be used by evaluators, saving them from attempting to gather it retrospectively.

Peacebuilding uses both formative evaluation (conducted in the midst of an ongoing project to identify progress and opportunities for improvement) and summative evaluation (conducted at the end of the project to understand the changes achieved). However, the vast majority are completed without the aid of a baseline assessment, which depicts the state of a situation prior to an intervention and enables a comparison to be made between then and now. This undermines an evaluation's ability to draw firm conclusions regarding the degree of change that has occurred. The lack of monitoring data and baselines both raise accuracy challenges to the average peacebuilding evaluation.

One of the reasons that baselines are commonly not conducted is that the question only arises when the evaluation report deadline looms on the horizon.¹⁷ Better practice would be for evaluations to be considered during the project design phase, where they can be adequately planned and budgeted for.

Finally, one increasingly sees technical assistance positions for monitoring and evaluation within peacebuilding agencies or development organisations. Often housed in institutional learning units, these positions are responsible for providing technical guidance to programmatic teams on design, monitoring and evaluation issues. In addition, many agencies have produced evaluation guidelines or standards that are of high quality and align to professional evaluation standards.

However, these standards are not consistently applied to the evaluations conducted within the organisation. This is due to a variety of reasons: standard organisational challenges of ensuring

¹⁷ Another reason is the confusion between baseline assessments and conflict analysis (see Church/Rogers 2006, 62 for further discussion).

agency-wide adoption of policies are one contributing factor, as is resistance to evaluation. Yet even strict adherence to the guidelines would not automatically support the main accountability and learning goals of evaluation (this latter point will be discussed in more depth in *section 5.4*).

4.

Peacebuilding Evaluation: What Makes It Different?

Many in the peacebuilding field operate from the assumption that peacebuilding evaluation differs from evaluation of other fields in that it is more complex and challenging (Church/Shouldice 2002). This assumption is mirrored within each peacebuilding sector. As a result, there has been a proliferation of sector-specific evaluation manuals in the last two years: peace education, security sector reform, transitional justice, peace mediation, rule of law, and so on. In many ways this is an achievement for the field of peacebuilding, because sector-specific evaluation guides could galvanise greater adoption and integration of evaluation. That said, it has also contributed to the idea that each sector (and peacebuilding in general) is fundamentally different – which has, for some, led to a barrier of entry. From the outside (before delving into the guides) those who have learned the basics of design, monitoring and evaluation are led to believe that these frameworks are irrelevant when it comes to peacebuilding.

Although this assumption of difference is reasonable, it also needs to be appropriate in its emphasis. At present, there is a tendency within the field to focus far too much on the unique elements of peacebuilding and not enough on the basic processes that enable adherence to core evaluation standards. What is needed is clarity in terms of where peacebuilding is different, so that peacebuilding evaluations can develop unique responses to those differences, whilst capitalising on good practice from the evaluation discipline where the difference is not relevant.

Three lenses have been used to reflect on where peacebuilding evaluation differs from other fields: process, content and context. Within each lens, peacebuilding presents notable differences as well as issues where the difference is a matter of degree.

In any evaluative *process* one must understand what issues or questions are to be explored: commonly called the evaluation criteria. Though there are no international rules as to what those should be, there is an increasing tendency to use the ones put forward by the OECD-DAC for all foreign assistance. The DAC posits six core criteria for all fields; peacebuilding included:

- relevance,
- effectiveness,
- efficiency,
- impact,
- sustainability,
- coherence.

The definitions of each of these are then tailored to the field in which they are to be applied – such as humanitarian aid or peacebuilding.

Though the criteria are not unique to peacebuilding, one does need to be knowledgeable about how they have been tailored to the field and the implications this has for the evaluation process. For instance, the OECD-DAC defines relevance for development as “the extent to which the aid activity is suited to the priorities and policies of the target group, recipient and donor”.¹⁸ For peacebuilding, the definition reads “the extent to which the objectives and activities of the intervention(s) respond to the needs of the peacebuilding process” (OECD-DAC 2008, 39). Though a seemingly small difference, the nuance is significant in terms of the work required by an evaluator. For a development evaluation, one would utilise a needs assessment, whereas for a peacebuilding evaluation one would need to draw from a conflict analysis.

Having to ground an evaluation in an analysis of the context again is not unique to peacebuilding. Context analysis is clearly stated as part of the accuracy standard while it also appears as a distinct step in some evaluation approaches. For instance, the CIPP model (context, input, process and product) developed by Daniel Stufflebeam (Stufflebeam et al. 2000) uses it to ascertain the quality of the programmatic goal. However, the peculiarities of conducting a conflict analysis (as distinct from a context analysis) do require a degree of peacebuilding expertise. This ensures that the evaluator is familiar with the common conflict analysis frameworks and understands the challenges in making linkages between the analysis and evaluative conclusions. One sees a similar challenge in evaluations in the development and humanitarian field that are assessing the conflict sensitivity of an intervention, as that is also based on a conflict analysis.

In addition, the OECD-DAC *Guidance on Evaluating Conflict Prevention and Peacebuilding Activities* recommends three optional criteria to be used: linkages, coverage and consistency with values/conflict sensitivity. Determining how an evaluation would adequately address these issues, especially linkages and consistency with values, remains a unique challenge to the field of peacebuilding evaluation.

The second lens, evaluation *content*, also has some common and some unique elements. In all evaluation processes, one needs to compare the findings with standards in order to ascertain the value of the achievement. For instance, that 65 percent of community leaders from across Kosovo who have finished a training programme on alternative dispute resolution are able to distinguish between negotiation and mediation is an example of a research finding. An evaluation then ascertains the value of this finding by comparing it to some form of standard. In other words, it assesses whether this was an excellent, good, average or poor result.

As peacebuilding and its sub-sectors do not have clear standards of comparison, quality evaluations must determine a way to draw conclusions of value. Here the difference to other fields diminishes as there are some classic processes that can be drawn from the evaluation discipline to create these standards, such as literature reviews, meta-evaluations and participatory processes.

Assessing value also applies to process – one assesses the quality of the process against process standards; for instance good practice in adult education might be used to ascertain the

18 See www.oecd.org/document/22/0,2340,en_2649_34435_2086550_1_1_1_1,00.html.

quality of a trainer. In this case, the difference again is one of degree, as there are many fields from which process standards can be drawn. They must, however, be adapted to both the cultural context and the realities of a conflict environment. Working with traumatised people or in environments where speaking openly is dangerous needs to be duly considered in the development of good process standards.

Peacebuilding requires some unique expertise when generating actionable recommendations specific to a peacebuilding practice, such as how to improve a mediation process or the quality of a conflict analysis. Yet this is not an absolute either, as it is possible to draw recommendations out through the data collection process. Also, there are areas commonly found in evaluations that do not require unique peacebuilding expertise, such as management, working with partners, communication or budgeting.

The third lens is the *context* in which the evaluation takes place, which raises numerous unique challenges for an evaluator. Significantly, conflict and post-conflict environments pose real challenges to data collection. As there are good books dedicated to the issue of research in conflict zones (e.g. Smyth/Robinson 2001), this discussion will flag only a few.

First, accessing key people to participate in the data collection process can be very difficult, especially when dealing with belligerents to the conflict, internally displaced people and refugees. All of these groups may be physically difficult to find or may not wish to participate due to uncertainty or fear.

Second, once key people have agreed to provide information there are also considerations to keep in mind when engaging with highly traumatised people. The data collection process should not re-traumatise the individual, which is a real concern where the topic is sensitive, such as with gender-based violence, missing persons or torture victims.

Ensuring that an appropriate degree of sensitivity to the context is reflected in the data collection methods is critically important. For instance, for many years peacebuilding evaluators in Burundi could not ask participants if they were Hutu or Tutsi, though this was a critical piece of information. In this case, evaluators got around the issue by having interviewees indicate which neighbourhood in Bujumbura they resided in, as the city was so segregated that this gave a reasonable indication of ethnicity.

Third, all evaluations, regardless of the context, should do no harm. What is distinct for peacebuilding is that if an evaluation unintentionally causes negative effects, the consequences could be far more severe and violent than in other contexts. For instance, consider an evaluation that included a focus group that discussed violence between two communities. If the conversation was not facilitated skilfully it could quickly degenerate into accusations and possibly (more) violence.

Similarly, evaluators must try to ensure the safety of participants in the evaluation at all times. Consider an assessment being conducted in Afghanistan of a programme that is trying to stop the flow of poppy-generated funding to militias and the Taliban. If the evaluators took written notes during interviews, which were subsequently obtained by either militia groups or the Taliban, the consequences for the interviewees could be dire.

Overall, understanding the unique aspects of peacebuilding is important so that evaluators can find custom responses. Equally important is to understand what is not unique to peacebuilding

evaluation, so that good practice in the evaluation field can be capitalised upon for the benefit of peacebuilding.

To date, peacebuilding has had difficulty making this distinction due to a lack of exposure to the evaluation profession or academic discipline. This is understandable, as the majority of monitoring and evaluation (M&E) staff within organisations have evolved into their role from field positions. Such learning on the job offers great richness in some ways, but it also commonly lacks the broad overview of the evaluation field or depth of knowledge beyond one or two approaches. This also holds true for many evaluators who are substance ‘experts’ rather than evaluation specialists. Yet expertise in evaluation is needed because it provides the framework from which the substantive information is then processed and organised.

One of the consequences of this lack of exposure to contemporary evaluation practice is a rather out-dated perception of evaluation. Basing perceptions on out-dated practice fuels the belief that evaluation of peacebuilding is entirely different or more complex. In reality, there are increasingly sophisticated and nuanced approaches that have been well tested. They could – with appropriate refinement – enormously improve the quality and, therefore, utility of evaluations in the peacebuilding field.

Finally, though, while much improvement could be derived from greater inclusion of good practice from the evaluation discipline, it is also important to remember that this field does not have a set of answers that will solve all peacebuilding evaluation dilemmas.

5. Open Questions and Challenges

Peacebuilding evaluation, regardless of the level one seeks to assess – outcomes, strategy or PWL – faces numerous challenges. These challenges can be broken down into methodological problems and issues at the heart of evaluative practice. This discussion focuses on the latter as they often underpin the former, methodological challenges. Furthermore, the issues at the heart of the evaluative practice are directly connected to the previously discussed “quality gap” in peacebuilding evaluation.

5.1

What is the Purpose of Peacebuilding Evaluation?

In theory, evaluation should never be an end in itself. It should be a mechanism that contributes to *accountability* and *learning* at a variety of levels: project team, office, organisation or peacebuilding field. Though this is reflected widely in the literature on peacebuilding evaluation, how it translates operationally is problematic.

This is particularly true for the role of accountability in peacebuilding and how evaluation specifically contributes to it. Again, principles are a useful starting point as one needs to have

a clear understanding of what accountability is in order to think about how it pertains to peacebuilding.¹⁹ There are two key elements to accountability: first, the provision of information to those to whom one is responsible, allowing them to assess the appropriateness of the information. The second element is the ability of those who receive the information to use it to judge and take action. The accountability relationship classically flows upwards to the donor, but can also be ‘downwards’ to programme participants or ‘horizontal’ to partners.

The second purpose of evaluation is to contribute to learning: how a team, agency or field adapts its practice based on a reasoned reflection of new information. Generally, learning occurs either in single- or double-loop form. Single-loop learning is based on reflection that focuses on those issues that fit within existing assumptions around strategy, such as how to improve a training programme. Double-loop learning questions the underpinning assumptions and theories behind the work, which can challenge the fundamentals upon which a peacebuilding programme is based (Pasteur 2006; Britton 2002). In order to catalyse single- or double-loop learning, an evaluation must know which of the two is desired from the outset so that the methodology can be developed accordingly.

There are different perspectives on how to interpret the *dual purposes of evaluation*. Some see them as complementary – each stronger because of the other – while others view them as contradictory forces. Those who think of them as contradictory view accountability as undermining learning because those being evaluated (the evaluand) will not speak candidly about weaknesses or failures. As a 2010 INTRAC paper points out, accountability can affect “the honesty and integrity with which information and analyses are presented” (Simister/Smith 2010, 7). This perspective maintains that learning needs a safe environment where the evaluand can identify lessons from successes and failures without fearing the consequences.

The alternative perspective sees the two mandates as complementary, whereby the act of being accountable catalyses action around learning as there are consequences if appropriate responses to an evaluation are not made. As One World Trust, a key player in spearheading accountability issues in civil society, noted in a recent project report, “the need to be open and transparent, engaging and responsive and continually learning [is] crucial to accountability” (Sastry 2009, 14).

In this perspective there is a clear distinction being made between financial accountability, compliance and performance accountability. Financial accountability refers to “answerability for the allocation, disbursement and utilisation of funds” (Molund/Schill 2004, 12); while compliance refers to reporting activities conducted against activities promised. Performance accountability, on the other hand, focuses on results and the process of providing pertinent information to key stakeholders to enable them to assess whether the agency has met the full range of their obligations. It is performance accountability that can be a catalyst to learning.

¹⁹ The meaning of accountability, in this discussion, has been drawn from the work of the Humanitarian Accountability Partnership (HAP) as well as the Overseas Development Institute. See Humanitarian Accountability Partnership, Frequently Asked Questions; available at www.hapinternational.org/other/faq.aspx and Davis 2007.

5.2

Is the Peacebuilding Field Learning from Evaluation?

Read any of the main guides to evaluation for peacebuilding and it becomes clear that the majority of the peacebuilding community subscribes, often implicitly, to the view that accountability and learning are contradictory. The field prioritises evaluation for learning and where possible moves away from accountability, outside of the tasks associated with obligatory upwards accountability, such as compliance and fiscal accountability (Tsadik 2008).

This choice, although commonly implicit, is noteworthy. Particularly so as the humanitarian field has tended to move in the opposite direction – towards accountability – as evidenced by the significant attention it has recently given to understanding what accountability is for humanitarian work, how it aligns with humanitarian principles and the role of evaluation in achieving accountability (Davis 2007). As the 2008 *Humanitarian Accountability Report* notes, “the overall impression gained is of a widening and deepening of accountability within the humanitarian system during 2008. [...] The sense is of accountability to intended beneficiaries and local communities becoming increasingly, if tentatively and somewhat patchily, integrated within the operational approach of a growing number of agencies” (Borton 2008, 64).

Given peacebuilders’ emphasis on evaluations contributing to learning, the question arises: is the field actually learning from the evaluative process and products? In other words, does the average evaluation report provide concrete input that enables single- or double-loop learning? It is difficult to speak for the field as a whole, as there certainly are some quality evaluation processes that have catalysed learning, however evidence suggests that this is not yet the norm. **This is particularly true at the level where a programme evaluation may have contributed to team learning, but is not doing so for the broader peacebuilding community.**

The reasons for this vary according to agency types, sizes and contexts; however, there are some clear commonalities across the field. Despite a commonly heard commitment to learning by many, most agencies do not have an organisational culture where learning is a central pillar (Hopp/Unger 2009). This is visible in the absence of supported structures, opportunities or time for staff to allocate to substantive learning. Yet if this is not supported by action it is difficult to translate lessons identified in an evaluation to lessons *learned*, as there are no set processes that enable this uptake.

Furthermore, there is the question of incentives and consequences. Outside individual commitment, where are the incentives within organisations to encourage staff to prioritise learning? Conversely there are few, if any, direct tangible consequences to organisations or individuals for *not* learning. Without this enabling context, it becomes very difficult to ensure that evaluations indeed contribute to learning.

A second common constraint is that many evaluations are not designed to support learning, despite the organisation’s stated commitment to this end. The intended ‘learners’ may not be clearly identified or may be too numerable to be realistic. The Terms of Reference may focus on interesting areas of inquiry that are simply not actionable, regardless of the findings. The report deadline may be weeks after a key learning opportunity, such as a new proposal or restructuring of the organisation. Learning also requires an evaluation that is deemed credible,

therefore, poor-quality evaluations, of which there are far more than high-quality ones, are not good material for producing new knowledge.

Moreover, a report identifying lessons is not the same as lessons learned. Thus a process must be planned to facilitate this transfer, which frequently gets omitted due to lack of time or simply because people are not aware of it. It is within these issues that lack of evaluation capacity can be seen, as most practitioners do not have enough experience to bring the goal of the evaluation in line with its subsequent actions.

Finally, learning is not isolated to the evaluation report, but can also occur throughout the evaluation itself. Called “process use”, it can offer as much learning as the final report when intentionally planned. Yet, process use is rarely fully capitalised upon in peacebuilding evaluation processes.

The operational challenges to catalysing learning are not unique to the peacebuilding field. Evaluation struggles to fully deliver in this way in many sectors. However, there are lessons and approaches that can be drawn from the evaluation discipline that would support greater peacebuilding learning profit from evaluations.

5.3

Does Evaluation Feed Accountability in Peacebuilding?

To respond to the question of whether or not evaluation is feeding accountability in peacebuilding, one first needs to have clarity on what this means in peacebuilding and on existing accountability mechanisms. However, accountability as a concept has not yet been tackled by the peacebuilding field, as evidenced by the paucity of specific material available. This is striking in its own right, but particularly when compared to the humanitarian community, who are in the midst of a dynamic debate on this issue, led in part by the Humanitarian Accountability Partnership (HAP).

Currently, accountability mechanisms in peacebuilding are almost exclusively upwards in nature. As Simon Fisher and Lada Zimina state, accountability “is often much more to funders and governments than to the people they work with and the communities they serve. Project proposals are, due to understandable practical constraints, often made with minimal consultation between local CSOs and INGOs [...]. This does not sit easily alongside a commitment to positive peace, justice and wellbeing of people and their communities” (2009, 12). Upwards accountability generally, though not always, entails some combination of compliance, fiscal accountability and an attempt at ascertaining impact.

Outside of upwards accountability, peacebuilding to date has shied away from engaging in the accountability discussion despite its picking up momentum in the larger international field. For instance – to name but a few initiatives in addition to the humanitarian efforts described earlier – One World Trust is pursuing projects developing the concept of accountability within civil society, the Hauser Centre for Non-Profit Organizations at Harvard University along with

CIVICUS is promoting the *INGO Accountability Charter*²⁰ and the International Council on Human Rights Policy is asking what accountability means for human rights (ICHRP [no date]). At the same time, one does not find accountability on the agenda of peacebuilding conferences, and there are hardly any articles on the topic in the academic literature. Accountability has remained nearly a non-issue in the peacebuilding field. As a result it is not well understood by practitioners operationally, nor well defined academically – particularly in terms of any nuance that peacebuilding itself may demand.

Reflecting on the meaning and mechanisms of accountability raises numerous important challenges relevant to power, politics and ownership for the peacebuilding field. Only a few will be covered in this section, to give a flavour of the dialogue needed.

One challenge pertains to whether or not accountability requires a **shared agenda** between those doing peacebuilding and those participating in the programming. If, according to our definition, accountability is the provision of information to those to whom one is responsible and if we can assume that this means communities in conflict; what happens when information about the agenda (goal) of the work is not to the liking of one or more groups within that community? There is a substantial amount of politics present in the agenda of peacebuilding, which almost by definition implies that something could be contested. Compare this to humanitarian work, where beneficiaries have a basic human need that requires assistance, such as clean water or emergency health care, which the humanitarian agency seeks to meet: though priorities may be contested, the overarching agenda – **meeting human need – is clear and shared.**

By extension this includes the question of who gets to define peace: the donor community, the local population, the warlords or the peacebuilding NGOs? As this is the overarching vision that programming is driving towards, it has a substantial influence on what work gets done. Whoever controls defining the vision has power. A classic illustration can be seen in the multiethnic agenda of the international community in Kosovo. There are many Kosovar Serbs and Albanians who would not subscribe to the vision of a multiethnic state, yet are participating in peacebuilding programmes. If the programme were directly accountable to these communities, what would the impact be on the peacebuilding work?

This in turn brings us to the challenge of who the peacebuilding field is responsible towards. To date, the field has largely turned ‘upwards’ to the donor community, but should the donor world be the only party involved in accountability relationships? **This is a complex challenge which raises numerous questions. For instance, are all communities – victims, militias, refugees, government representatives – equal in terms of accountability relationships for those involved in peacebuilding?**

Another phrase in the definition of accountability, “the provision of information”, raises further challenges: the potential for misuse of information. In highly volatile contexts, misused information could spark unrest or lead to local partners being threatened. Although this is a legitimate concern and needs to be considered seriously, there is also a need for the field to exercise self-discipline in assessing where information could lead to security threats. Otherwise,

20 See www.ingoaccountabilitycharter.org/.

the argument has the potential to be used as a blanket block to information, even in relatively secure situations. For instance, information contained in an evaluation of mediation training or a public awareness campaign on the role of police in a post-conflict society is not generally so sensitive that it would endanger local partners.

Another way of considering this issue is from the perspective of who has power and control. Accountability involves shifting control of information to actors outside the organisation. Furthermore, it gives those actors power of response, which is key to the accountability relationship. Compare this with evaluations for the purpose of learning as they are commonly implemented: the information is generally kept within the organisation; where actors internally decide if and how it will be acted upon and who else will receive it. In considering whether to share information, such as the contents of an evaluation, with communities, it must be remembered that there is a difference between protecting partners and protecting one's own power and control.

The impulse for a wider accountability discussion should come from peacebuilding agencies themselves rather than waiting for a watchdog group or aggrieved citizens in a conflict zone to demand it. As long as the concept of accountability for peacebuilding is not clear, it will remain difficult for the field to develop evaluation systems that contribute to it.

5.4

Evaluation as a Means to Different Ends: Bad Practice

So far, this section has covered a variety of explanations as to why evaluation is not yet fully delivering on learning and accountability for peacebuilding. There is another important impediment that deserves attention: the use of evaluation for other ends.

Far too often, instead of accountability and/or learning, there are various other purposes behind evaluation: public relations, fundraising and the justification of existing programming or predetermined decisions. This may be explicitly stated in the Terms of Reference; for example when a multilateral agency cites the need to raise additional funds from member states as the rationale behind the evaluation. More commonly, the stated purpose is learning and/or accountability, while the actual orientation only becomes apparent by reading between the lines: the types of questions in the Terms of Reference, resources allocated to the process, guidance given by the evaluation manager, or the type of feedback received on the draft report. For instance, adjectives such as 'very' or 'extremely' may be generally acceptable if applied to statements of success, but scrutinised in detail if pointing towards issues of under-performance.

There are numerous explanations as to why accountability and learning are not driving the evaluation agenda in peacebuilding. The insecurity within implementing agencies caused by the aid 'system' is a primary culprit. Funding that is offered on 12, 18 or 24 month intervals diminishes the ability to develop productive, trust-based partnerships with donors. This, coupled with the common perception amongst implementing agencies that there is only room for success and not for failures, leaves many with the conviction that identification of poor performance through an evaluation will have dire consequences in terms of funding or organisational reputation. This in turn "induces a culture of caution: only successes are reported

in any detail, though failures are inevitably frequent and are often the most fertile arena for learning” (Fisher/Zimina 2009, 24).

Many donors are aware of – and frustrated with – this success-only bias. For instance, in 2009 one major peacebuilding donor reviewed all peacebuilding evaluations done through their department and was surprised to discover they had a 100% success rate in programming.

One cannot only blame the overarching system, however, as individuals also play a key role in this dynamic. In donor and implementing agencies alike there is a complex set of personal incentives and disincentives to use evaluation for accountability and learning. Why would a desk officer with aspirations to be head of the regional office want to report that the project he granted several million dollars did not create change on the ground? Moreover, what consequences are there for this individual for *not* exploring thoroughly whether the supported programming was effective? In reality, there are few agencies that have put in place incentives for individuals to adopt evaluation for promoting learning and accountability. Given the lack of information that participants in peacebuilding programming generally receive from organisations, coupled with their lack of power (see *section 5.3*), there are also few consequences.

Finally, despite evaluation having been firmly on the peacebuilding agenda for a decade, the average peacebuilding team and agency lacks the knowledge and skills necessary to establish and manage quality evaluation processes. For instance, practitioners are rarely fluent in the core decisions pertinent to evaluation that need to be made during project design (such as which evaluation criteria to choose or the value of a baseline). This is often even the case when organisations have guidelines on evaluation because the average staff person is not conversant enough in their implementation to be able to apply them.

Although the capacity gap is a reality, this argument is frequently overused and needs to be challenged appropriately. Some of the most common distortions to evaluation do not stem from a lack of understanding of complicated evaluation applications. **Instead they are the result of an orientation that perceives evaluation as a means to a fundraising or public relations end, where there are few boundaries on what means are acceptable.** For instance, when an evaluation is conducted for the purpose of fundraising or public relations it is common for the evaluator to be ‘requested’ to remove paragraphs that are seen to be critical of the programming, to confidentially write an internal and a public report,²¹ or only to engage with select audiences.

The underlying values that make these actions unacceptable – fair representation and honest reporting – are not unique to evaluation; they are core values of peacebuilding as well.²² **Cynthia Cohen**, a scholar researching the role of ethics in conflict and peacebuilding, states that values are “central to all work in the fields of coexistence and reconciliation” (2001, 1.8). Research indicates that the values of the field do coalesce into some commonly held principles. Anderson and Olson’s *Confronting War* found that “surprisingly among a range of peace practitioners engaged in a wide variety of activities, there is a great deal of agreement about the elements of

21 It is not always unacceptable to do this in practice. The public report should then clearly state that it has been tailored to this audience and is not a comprehensive representation of the evaluation.

22 For the purpose of this chapter, principles, ethics and values are being used interchangeably for simplicity’s sake, though it is recognised that technically they are quite distinct concepts.

the ethics of peace practice” (2003, 27).²³ It is these ethics that need also to be used in managing or being subject to evaluation.

For instance, one of *Confronting War*’s findings was that peacebuilding practitioners believe that the end does not justify all means. This would make some of the actions described earlier incongruent with peacebuilding values. Another shared value found by *Confronting War* is that peace practice is honest, which includes individual accountability, sincerity and agency transparency. Further, practitioners also felt that honesty included being honest “about mistakes and mixed outcomes” (Anderson/Olson 2003, 29). If this is the case, each decision in the evaluative process should be grounded in an explicit commitment to honesty. This would stop many of the practices commonly attributed to a lack of evaluation capacity.

Yet currently, it is still “easy to throw away our values when it comes to evaluation”, as a colleague at the Fletcher School of Law and Diplomacy summed it up recently after reflecting on evaluations that he was familiar with.²⁴

6.

Recommendations: What Can Researchers, Practitioners and Donors Do?

The three main international actors in peacebuilding – academia/researchers, implementing agencies (both INGOs and CSOs) and donors – all have a role to play in advancing the quality of peacebuilding evaluation.

First and foremost, a *field-wide conversation* needs to occur on the meaning and operational implications of accountability within peacebuilding. Networks such as the Washington-based Alliance for Peacebuilding or the Global Partnership for the Prevention of Armed Conflict (GPPAC) would be ideal candidates to coordinate and drive this process from behind. Critical to the success of this conversation would be to ensure that space is provided for those living in conflict or post-conflict societies to convey their experience with peacebuilding agencies and voice their perspective on what accountability should mean.

Though the conversation should start at the conceptual level, it should not lose track of operational realities and should ultimately seek to make a *meaningful contribution to peacebuilding agencies’ accountability processes*. This would, of course, include the role evaluation should play. There are processes currently underway that one could draw from in order to understand how best to engage in this topic. Examples include the work of One World Trust on accountability in research, or the new effort by the International Council on Human Rights Policy to understand human rights principles and NGO accountability. As with accountability

²³ What Anderson and Olson found is similar to International Alert’s 1998 *Code of Conduct for Conflict Transformation Work*, which references ten principles such as the primacy of people in transforming conflict, impartiality, independence and accountability.

²⁴ Personal conversation, November 2008.

work in the humanitarian community, this would not be a finite project, but rather an ongoing endeavour to consistently improve the way the field honours its accountability obligations.

As accountability and learning overlap in their role as motivations behind evaluation, improvements on the accountability side would have implications on the learning agenda. An important piece of work to be done by evaluation specialists or researchers would be to tease out the ramifications for learning that a greater emphasis on performance accountability would have.

In terms of learning, those commissioning evaluations need to pay greater attention to *ensuring that learning is embodied throughout the evaluation process*. Here the evaluation discipline has a lot of practical experience to offer the peacebuilding field so that evaluations can deliver on organisations' learning expectations. At the field level, donors have a specific role to play in enabling learning. If the donor community consistently required evaluations to be made public, this would make possible more evaluation synthesis and meta-evaluation.²⁵ Both processes offer learning at a more general or overarching level and can be organised by evaluators, donors or researchers.

To address the quality gap in peacebuilding evaluation, a first important step would be to *initiate greater overlap between those conducting peacebuilding evaluations and those in the evaluation discipline*. This could be encouraged in numerous ways, for instance by adding evaluation expertise to the recruitment qualifications of prospective evaluators, combining practical experience with knowledge of best practice in evaluation. Possible indicators of expertise can be having currency in evaluation literature, demonstrating ongoing professional development in evaluation or presenting work at national evaluation conferences.

Ensuring the application of agency or donor standards in evaluation processes would make a significant difference in terms of quality. Donors could play a greater role in this process as well, by reviewing any prospective evaluation against the four main evaluation standards of *utility, feasibility, propriety and accuracy*. Further, they could model good evaluation practice by ensuring that they followed their own institution's evaluation guidelines.

Also, *greater emphasis needs to be given to evaluation during the project design stage* by implementing agencies. Decisions on who the evaluation will inform, when it will be done and what it will explore, have an impact on budget and work planning choices and thus need to be considered before a proposal is submitted. Developing a real-cost budget or using the industry guidance of adding 5-10 percent of project costs will provide adequate resources to investigate questions of significance that can in turn generate single- and double-loop learning. Furthermore, thinking through the evaluation process at this stage creates the opportunity to involve an evaluator in developing a monitoring system and conducting a baseline *prior* to the project start.

Finally, *more investigation is needed into key methodological challenges*, such as how one ascertains relevance, the comparative advantage of different theories of change, and dealing with multiple possible reasons as to why change may have occurred (*the so-called attribution-contribution challenge*). Having better approaches to these issues would also contribute to evaluation's ability to deliver on its learning and accountability promises.

²⁵ Evaluation synthesis reviews a set of evaluations within a specific theme to identify the lessons for that thematic area; while a meta-evaluation does the same for evaluation practice.

7. Conclusion

Evaluation is increasingly seen as a mainstay of the field; with a greater number of dedicated positions within agencies, more evaluations and a burgeoning pool of peacebuilding evaluators. Research has expanded areas of inquiry beyond outcomes-focused evaluation to look at the contribution that programmes make to strategy and peace writ large. Despite all this, evaluations are not yet contributing to learning and accountability in peacebuilding to their full potential.

While methodological shortcomings definitely play a role, so, too, do two other key phenomena. The emphasis on how different peacebuilding (and its evaluation) is from other fields has resulted in an evaluation practice that is not grounded in the good practice elaborated in the evaluation discipline. This is most overtly seen in the frequency with which peacebuilding evaluations do not meet the standards put forth by the professional evaluation community. It also plays an important role in the tendency to hire peacebuilding substance experts rather than those with expertise in evaluation.

Underpinning this general lack of adherence to quality standards is a more fundamental issue: the distortion of the purpose of evaluation away from learning and accountability and towards fundraising, public relations and programme justification. The way this plays out in practice is a direct challenge to the espoused peacebuilding values of honesty and of the ends never justifying the means.

There is certainly room for improvement in the technical applications of evaluation. Yet the area most in need of attention is the alignment of peacebuilding values in the implementation of evaluations. With this will come greater attention to the meaning of accountability and more incentives for quality evaluation practice, which in turn will drive the appropriate adoptions, adaptations and innovations needed in the technical realm.

8. References

- Anderson, Mary B. and Lara Olson 2003. *Confronting War: Critical Lessons for Peace Practitioners*. Cambridge, MA: The Collaborative for Development Action, Inc.
- Borton, John 2008. An Overview of Humanitarian Accountability in 2008, in: Humanitarian Accountability Partnership (ed.). *The 2008 Humanitarian Accountability Report*. Geneva: The Humanitarian Accountability Partnership International, 7-70.
- Britton, Bruce 2002. *Learning for Change: Principles and Practices of Learning Organisations*. Sundbyberg: Swedish Mission Council. Available at [http://portals.wi.wur.nl/files/docs/ppme/learningforchange\[1\].pdf](http://portals.wi.wur.nl/files/docs/ppme/learningforchange[1].pdf).

- Brusset, Emery and Ralf Otto 2004. *Evaluation of Nashe Maalo: Design, Implementation, and Outcomes – Social Transformation Through the Media, On Behalf of Search for Common Ground*. Ohain: Channel Research.
- Church, Cheyanne 2008. Peacebuilding Evaluation: From Infancy to Teenager [Introduction], in: *New Routes. A Journal of Peace Research and Action*, 13(3), 3-7.
- Church, Cheyanne and Mark M. Rogers 2006. *Designing for Results: Integrating Monitoring and Evaluation in Conflict Transformation Programs*. Washington: Search for Common Ground.
- Church, Cheyanne and Julie Shouldice 2003. *The Evaluation of Conflict Resolution Interventions: Framing the State of Play, Part II: Emerging Practice & Theory*. Derry/Londonderry: INCORE International Conflict Research.
- Church, Cheyanne and Julie Shouldice 2002. *The Evaluation of Conflict Resolution Interventions: Framing the State of Play*. Derry/Londonderry: INCORE International Conflict Research.
- Cohen, Cynthia 2001. *Working with Integrity: A Guidebook for Peacebuilders Asking Ethical Questions*. Waltham: The International Center for Ethics, Justice and Public Life.
- Davis, Austen 2007. *Concerning Accountability of Humanitarian Action*. London: Human Practice Network, Overseas Development Institute. Available at www.odihpn.org/documents/networkpaper058.pdf.
- Fisher, Simon and Lada Zimina 2009. Just Wasting our Time? Provocative Thoughts for Peacebuilders, in: Beatrix Schmelzle and Martina Fischer (eds.) 2009. *Peacebuilding at a Crossroads? Dilemmas and Paths for Another Generation*. (Berghof Handbook Dialogue No 7.) Berlin: Berghof Research Centre, 11-35. Available at www.berghof-handbook.net/documents/publications/dialogue7_fishzim_lead.pdf.
- Hopp, Ulrike and Barbara Unger 2009. Time to Learn. Expanding Organisational Capacities in Conflict Settings, in: Beatrix Schmelzle and Martina Fischer (eds.). *Peacebuilding at a Crossroads? Dilemmas and Paths for Another Generation*. (Berghof Handbook Dialogue No 7.) Berlin: Berghof Research Centre, 77-86. Available at www.berghof-handbook.net/documents/publications/dialogue7_hoppunger_comm.pdf.
- Goodhand, Jonathan and Paul Bergne 2004. *Evaluation of the Conflict Prevention Pools, Afghanistan*. London: Bradford University, Channel Research Ltd, PARC & Associated Consultants.
- INCORE & IDRC 2009. *Evaluating Research in and on Violently Divided Societies, A Background Note*. Derry/Londonderry: INCORE & IDRC.
- ICHRP [no date]. *ICHRP Online Discussion Forum on Human Rights Principles and NGO Accountability: An Approach Paper*. Geneva: International Council on Human Rights Policy (ICHRP). Available at www.ichrp.org/approach_paper.pdf.
- International Alert 1998. *Code of Conduct for Conflict Transformation Work*. London: International Alert. Available at www.international-alert.org/pdf/coc_full.pdf.
- Molund, Stefan and Göran Schill 2004. *Looking Back, Moving Forward. Sida Evaluation Manual*. Stockholm: Edita.
- OECD-DAC 2008. *Guidance on Evaluating Conflict Prevention and Peacebuilding Activities, Working Draft for Application Period*. Paris: OECD Publishing. Available at www.oecd.org/secure/pdfDocument/0,2834,en_21571361_34047972_39774574_1_1_1_1,00.pdf.

- Pasteur, Katherine 2006. Learning for Development, in: Rosalind Eyben (ed.). *Relationships for Aid*. London: Earthscan. Available at www.livelihoods.org/lessons/Learning/OrgLearn.html.
- Patton, Michael Quinn 2008. *Utilization-Focused Evaluation*. London: Sage Publications.
- Rossi, Peter Henry, Mark W. Lipsey and Howard E. Freeman 1999. *Evaluation: A Systematic Approach*. 6th edition. London: Sage Publications.
- Royal Norwegian Ministry of Foreign Affairs 2004. *Towards a Strategic Framework for Peacebuilding: Getting Their Act Together*. Overview report of the Joint Utstein Study of Peacebuilding. Available at www.regjeringen.no/upload/kilde/ud/rap/2000/0265/ddd/pdfv/210673-rapp104.pdf.
- Sastry, Deepti 2009. Civil Society Accountability: Reflections and the Ways Forward, in: *Civil Society Voices. Special Issue on Good Governance*, Voluntary Action Network India, June 2009, 12-14. Available at vaniindia.org/images/CSVJune2009%28English%29.pdf.
- Simister, Nigel and Rachel Smith 2010. *Monitoring and Evaluating Capacity Building: Is It Really That Difficult?* Praxis Paper 23. Oxford: INTRAC. Available at www.intrac.org/data/files/resources/677/Praxis-Paper-23-Monitoring-and-Evaluating-Capacity-Building-is-it-really-that-difficult.pdf.
- Smyth, Marie and Gillian Robinson 2001. *Researching Violently Divided Societies: Ethical and Methodological Issues*. New York: United Nations University Press.
- Spencer, Tanya 1998. *A Synthesis of Evaluations of Peacebuilding Activities Undertaken by Humanitarian Agencies and Conflict Resolution Organisations*. London: Active Learning Network on Accountability & Performance in Humanitarian Action (ALNAP). Available at www.alnap.org/pool/files/tspace.pdf.
- Stufflebeam, Daniel L., George F. Madaus and Thomas Kellaghan (eds.) 2000. *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*. 2nd edition. Boston: Boston Springer Science & Business Media.
- Tsadik, Hannah 2008. May I Ask, Where Is My Reconciled Heart? Overcoming the Accountability Imbalance, in: *New Routes. A Journal of Peace Research and Action*, 13(3), 7-11.

Recommended Resources

American Evaluation Association: www.eval.org
 Canadian Evaluation Society: www.evaluationcanada.ca
 European Evaluation Society: www.europeanevaluation.org
 International Development Evaluation Association: www.ideas-int.org
 Monitoring and Evaluation News: <http://mande.co.uk>

[All websites accessed 26 August 2010.]