



*practical learning for
effective international action*

Working Papers on Program Review & Evaluation: #4

**An Alternative to Formal Evaluation of
Peacebuilding:
Program Quality Assessment**

Cordula Reimann, Diana Chigas & Peter Woodrow

Reflecting on Peace Practice Program

October 2012

CONTENTS

About CDA Collaborative Learning Projects	i
About the Working Papers on Program Review and Evaluation	ii
I. Introduction	1
II. General Remarks on the Need for Use of Evaluations in Peacebuilding	2
III. Peacebuilding Program Quality Assessment	6
IV. Using RPP Concepts in Peacebuilding Program Quality Assessments	10
V. Conclusion	19
References	20
Annex: Peacebuilding Program Quality Assessment tool	

ABOUT CDA COLLABORATIVE LEARNING PROJECTS

OUR VISION

CDA strives for a world in which people are supported to enhance their resilience, drive their own development and resolve conflicts without resorting to armed violence.

OUR MISSION

CDA facilitates collaborative learning processes among humanitarian, peace, development and corporate practitioners and policy-makers to ensure that their support is effective and accountable.

OUR PROGRAMS

The oldest of CDA's programs, **DO NO HARM (DNH)**, seeks to identify the ways in which international humanitarian and/or development assistance given in conflict settings may be provided in ways that avoid exacerbating the conflict and maximize potential contributions to peace processes. DNH's conflict sensitivity tools are now used by practitioners around the world to design and implement assistance efforts that improve the outcomes for affected communities.

The **REFLECTING ON PEACE PRACTICE PROGRAM (RPP)** is an experience-based learning process involving agencies whose programs attempt to prevent or mitigate violent conflict directly. One of the central insights from the RPP case studies is that, while individual peacebuilding efforts may be beneficial, they do not necessarily contribute to – or, in the aggregate, “add up” to – what RPP calls Peace Writ Large. The RPP aims to improve the effectiveness of peace efforts and their positive role in building durable peace.

The **LISTENING PROGRAM (LP)** is a comprehensive and systematic exploration of the perspectives and ideas of people who live in societies that have been on the recipient side of international assistance, including humanitarian aid, peacebuilding activities, or human rights work. Listening and responding to local perspectives is critical for donors and aid agencies to live up to the global commitments they have made to listen to and involve local stakeholders and to improve the quality, effectiveness and accountability of their efforts.

The **CORPORATE ENGAGEMENT PROGRAM (CEP)** collaborates with companies to ensure that they have positive rather than negative impacts on the communities where they operate. As part of this approach, the CEP assists companies in the development and implementation of practical options to build constructive relationships with the communities where they work. Ultimately, CEP supports the contribution of the private sector to stability and sustainable development.

ABOUT THE WORKING PAPERS ON PROGRAM REVIEW AND EVALUATION

This Working Paper series is based upon findings regarding program review and evaluation processes from across all of the programs of CDA Collaborative Learning Projects (CDA). These include the Do No Harm, Reflecting on Peace Practice, Listening, and Corporate Engagement Programs.

All of CDA's programs are concerned with issues of impact, results, assessment and learning—as well as mechanisms for obtaining feedback and being accountable for program effects. This CDA Working Paper series aims to share reflections and lessons gleaned from across these programs for agencies who are undertaking their own program reviews and evaluations.

CDA produced these papers to offer practical support for practitioners seeking to evaluate the effectiveness and contributions of specific programs across several sectors. The papers do not represent final products and are working documents. CDA would like to acknowledge the generosity of the individuals and agencies involved in donating their time and experience and for their willingness to share their experiences and insights to generate and reflect on the learnings shared in these papers.

About the Authors

Cordula Reimann has worked for over fifteen years as a consultant, facilitator, trainer, researcher and lecturer in peacebuilding and conflict and peace studies. She has worked for international and Swiss governmental agencies and non-governmental organizations, including eight years at the Centre for Peacebuilding (KOFF) and at the Swiss foundation swisspeace, where she was head of analysis and impact of peacebuilding. Cordula conducted and led various trainings and evaluations on gender and peacebuilding, conflict sensitivity and the effectiveness and the impact of peacebuilding programs. Her field experience is mainly in South (East) Asia, Sub-Saharan Africa, and the Middle East. Her main areas of expertise are conflict sensitivity, strategic conflict analysis, impact assessment, gender, conflict and conflict transformation. She has published widely on gender, conflict and peacebuilding and conflict transformation theory. Cordula recently set up her own consultancy, coaching and training business called “core. consultancy & training in conflict transformation.”

Diana Chigas joined CDA in 2003 as Co-Director of the Reflecting on Peace Practice Project (RPP). She is also Professor of the Practice of Negotiation and Conflict Resolution at the Fletcher School of Law and Diplomacy. Prior to joining CDA, Diana worked as a facilitator, trainer and consultant in negotiation, dialogue and conflict resolution. Her work included development of strategies, training and advice on preventive diplomacy in the OSCE, and facilitation of dialogue in El Salvador, South Africa, Ecuador and Peru, Georgia/South Ossetia, and Cyprus. Diana

became involved with RPP as a participant in the first phase, as CMG's programs in Cyprus and Georgia/South Ossetia were the subject of case studies.

Peter Woodrow joined CDA as Co-Director of the Reflecting on Peace Practice Program (RPP). Peter is an experienced mediator, facilitator, trainer, and consultant. He formerly worked for CDR Associates in Boulder, Colorado where he mediated and facilitated issue resolution within organizations, as well as multiparty environmental and public policy disputes. He has also developed and implemented international programs in peacebuilding, consensus building, problem solving, decision making and inter-ethnic conflict resolution in Asia, Africa and Eastern Europe. Peter served on the Advisory Group during the first phase of RPP, before becoming co-director.

An Alternative to Formal Evaluation of Peacebuilding: Program Quality Assessment

I. Introduction

Since September 2003, CDA has been working with active peace programs to test the lessons of *Confronting War* as applied in practice. In this phase of RPP's work, practitioners have been concerned with how to measure the impacts of peacebuilding efforts. In this context, RPP is exploring how its findings concerning *peacebuilding* effectiveness and impacts can be integrated with good evaluation practice.

In her review of peacebuilding evaluations, Cheyanne Church concludes that the average peacebuilding evaluation would not be considered a high quality evaluation process according to professional standards in the evaluation field (Church 2011b, 6). And it seems safe to say that the peacebuilding community has been struggling for some years to find appropriate approaches and methods for evaluating the impact, effectiveness and relevance of their projects and interventions. While advances have been made by the OECD-DAC,¹ there is still no widely accepted methodological agreement about how best to conduct evaluations in complex conflict and peacebuilding contexts (Paffenholz 2011).

CDA Collaborative Learning Projects (CDA) and its Reflecting on Peace Practice Program (RPP) aim to help fill this gap. In 2009, RPP began a collaborative learning effort to develop practical support for practitioners seeking to assess the contributions of peacebuilding programs to Peace Writ Large. In a report commissioned by CDA, Church stresses that RPP can build upon the existing application of RPP concepts and findings in evaluation to support and strengthen the quality of peacebuilding program design. Improving program design, Church asserts, is a crucial first step toward effective evaluation (Church 2011b, 17-18).² As RPP tools and concepts already have significant acceptance among practitioners in the area of program design, they can be useful to strengthen evaluative thinking in peacebuilding.

Church also points to the need for an alternative to evaluation in instances of high resistance to evaluation or where the main focus is learning rather than accountability (see Church 2011b, 19). Additionally, some implementing agencies may lack sufficient funds to conduct a formal evaluation. This then suggests the need for an evaluative process geared towards facilitating collaborative learning among program teams, partners, and (where possible) donors, and

¹ OECD-DAC, *Evaluating Peacebuilding Activities in Settings of Conflict and Fragility—Improving Learning for Results* (Paris: OECD-DAC, 2012) (hereinafter “DAC Guidance”). Provisional guidance on evaluation of peacebuilding programs and activities was previously published in 2008 under the title, *Guidance on Evaluating Conflict Prevention and Peacebuilding Activities: Working draft for the application period* (Paris: OECD, 2008). This has been superseded by the 2012 DAC Guidance.

² Church's discussion in this context refers to the evaluability of peacebuilding programs. For guidance on how RPP might be used to guide evaluability assessments in peacebuilding, please refer to the third paper in this CDA paper series, Working Papers on Program Review & Evaluation, titled *Evaluability Assessments in Peacebuilding Programming*, also by Cordula Reimann, available at www.cdacollaborative.org. Last accessed December 5, 2012.

improving program design and implementation without attempting to meet the standards of formal evaluation.

Drawing on the findings of RPP regarding factors that contribute to effectiveness and impacts of peacebuilding, and on the experience of practitioners, this Working Paper offers an approach to reviewing program quality that addresses organizational needs for feedback, assessment and learning short of a formal evaluation. This Working Paper served as the basis for initial feedback and analysis by a mixed group of evaluation professionals and peacebuilding practitioners in December 2011. It has been revised based on that feedback and is now being disseminated more broadly.

In Section 2, this Working Paper offers some general introductory comments on evaluation in peacebuilding. Section 3 introduces the important features of program quality assessment in other fields and relevant insights from the latest thinking in the evaluation world. Against this background, Section 4 discusses the implications for the development of a “peacebuilding program quality assessment” process and presents a format based on RPP tools and concepts. The focus will be on promising entry points and methodological challenges in improving the quality of program design in peacebuilding and intervention projects. Finally, Annex A provides one tool for assessing program quality using RPP-based standards.

The paper is inspired by the standards and latest approaches in the evaluation field, a review of peacebuilding evaluations, the author’s experiences, the experiences of Swiss and international NGOs in evaluation, and assessments of conflict sensitivity and peacebuilding programs in different conflict contexts.

II. General Remarks on the Need for and Use of Evaluations in Peacebuilding

2.1 WHAT IS PEACEBUILDING EVALUATION?

In an article exploring the uses of evaluation in the peacebuilding field, Church draws on the many conceptualizations in the evaluation literature to offer a definition of peacebuilding evaluation as:

“...the use of social science data collection methods to investigate the *quality* and *value* of interventions which seek to stop violence from re-igniting or promote a positive change in the conflict context” (Church 2008, 3).

Church goes on to say that “quality” refers to “the manner in which the program is conceived and implemented,” while “value” involves an assessment of whether “the intervention effected a change that the participant communities identified as significant vis-à-vis the conflict” (Church 2008, 3). The focus on value requires making a judgment regarding the outcomes of a program, based on reliable evidence. Church and others conclude that while some peacebuilding evaluations manage to address quality, most fail to provide clear evidence regarding value

(Church 2008, 2011a).

2.2 OBSTACLES AND CHALLENGES TO EVALUATION IN THE PEACEBUILDING FIELD

To date, the peacebuilding field has been reluctant to accept and adopt the systematic use of evaluations, for a number of reasons.

First, peacebuilding evaluations confront unique challenges. For example, how does one measure the impact of a micro-level project on a macro-level conflict? There is the challenge of attribution, of course, but also of the highly complex non-linear processes of social change that cannot be captured by linear cause-effect logic.

Second, an established “evaluation culture” does not yet exist in the peacebuilding field. This is partly due to a certain ignorance in the peacebuilding community of the field of evaluation and its professional standards and practices. Others suggest more general reasons relating to the on-going need to establish peacebuilding itself as a legitimate field. Over the last ten to fifteen years, competition among a fast-growing number of peacebuilding organizations competing for scarce funding from governmental and private donors has hindered sharing and discouraged self-critical discussion. This, in turn, has limited field-wide learning from peacebuilding’s challenges and failures. The peacebuilding community, which has long fought for recognition within the well-established field of international relations, wants to protect the field’s reputation as a rigorous and professional community of practice (see Kawano-Chiu 2011, 8-9 and 34-35).

The twin pressures of scarce funding and the need to establish peacebuilding as a legitimate field have also led to a tendency among peacebuilding programs to ‘over-claim’ results. The Alliance for Peacebuilding depicts this as a vicious cycle of policy pressure and over-claiming by both donors and NGOs, leading to (the perception of) failing or ineffective peacebuilding programs. RPP Co-Director Peter Woodrow speaks of an “inadvertent and unrealistic collusion between funders and program implementers in making unrealistic claims” (Kawano-Chiu 2011, 19).

Examples of over-claiming of peacebuilding programs include lofty, unrealistic program goals, such as “harmony and co-existence between the two main ethnic groups.” Such vague and unrealistic goals have been shown to hinder effective project or program implementation (RPP 2010). Such goals are also hard to measure, and thus make it difficult to evaluate the program or its impact later on. At the same time, the pressure to “sell the success” of a particular peacebuilding intervention or program to funders and donors is strong. As a result, low quality evaluations whose main objective is to please or convince a donor are produced, and (often) not shared. This, in turn, results in a lack of transparency regarding failures and lessons that may have been identified.

Third, evaluators in peacebuilding often accept limited, if any, accountability and responsibility for use of the evaluation results. Evaluation reports are written, and local staff members are left

to their own devices, with little continuing, external support to implement the recommendations of an evaluation (*see also* Action Asia 2011, 16).

On a more positive note, as the Alliance for Peacebuilding report signals, there is increased interest in the peacebuilding community in exploring a range of “new” approaches in evaluation, such as the “most significant change” technique (Davies and Dart 2005), empowerment evaluation (Fettermann 2001), and developmental evaluation (Patton 2011), to name just a few. At the same time, fields related to peacebuilding, such as human rights and development, have made significant steps in their own evaluation practices. One inspiring and recent initiative is the “Admitting Failure” project launched by a Canadian NGO a few years ago.³ The project aims to “to establish new levels of transparency, collaboration, and innovation” by publishing project and program failures. While it remains to be seen how far admitting to failures results in meaningful learning from them, this inspiring and, in a way, liberating way of thinking about practice is still missing in the peacebuilding field.⁴

2.3 RATIONALE FOR ALTERNATIVES TO FORMAL EVALUATION

Some peacebuilding programs would not benefit from a formal evaluation. Some of the multiple reasons may be obvious and straightforward, while others are implicit and hidden.

- In some instances, an evaluation may merely be window-dressing, conducted to please the donor and its policies and written in an overly positive, polished tone. It might be designed mainly to serve as a form of compliance.
- Staff members may have reservations or resistance (open or hidden) to evaluation (Church 2011b, 19). Often, donors or commissioning agencies do not make the rationale for evaluation clear and transparent, and evaluated organizations and staff fear that future funding decisions are riding on the results of an evaluation.
- The local and international staff may perceive an evaluation as a punishment. In many cases, the staff fears failure, but would benefit from an open, non-judgmental space for learning from the project experiences, including their failures.
- The donor may not be asking for a formal evaluation, as there is either no budget for evaluation (or insufficient budget to conduct a proper evaluation), or there is a “donor fatigue” for commissioning evaluations. Yet the program team and the donor could benefit from some kind of evaluative review of the program.
- In some conflict and peacebuilding contexts, such as certain provinces in Afghanistan, Somalia, Iraq or Pakistan, there are heavy security constraints, and mistrust is very high. In these settings, a full-fledged evaluation might do harm and endanger the stakeholders involved (Action Asia 2011, 5). In these and other instances, traditional evaluation

³ “Admitting Failure” was created by the NGO Engineers Without Borders Canada and intended as a collaboration between like-minded NGOs, governments, donors and the private sector. They publish a yearly failure report. For the 2011 report, see <http://legacy.ewb.ca/mainsite/pages/whowere/accountable/FailureReport2011.pdf>. See also http://www.ted.com/talks/david_damberger_what_happens_when_an_ngo_admits_failure.html. Both links last accessed December 5, 2012.

⁴ Blum (2011) argues for whole-of-field solutions, lately pursued by some leading international peacebuilding organizations.

standards and expectations of rigor and validity in the data collection process are difficult to meet.

- Finally, overall, as Church notes, the average peacebuilding evaluation would not qualify as a high quality piece of work according to the five standards of quality in the evaluation profession: utility, feasibility, propriety, accuracy, and evaluation accountability (Church 2011a, 2011b). A low-quality evaluation may not be useful and could be misleading.

Despite these challenges and questions regarding quality, all peacebuilding programs aim at designing and implementing effective programs or project activities, and many desire and would benefit from processes grounded in *evaluative thinking*.⁵ The Illustrative Spectrum of Evaluative Processes (Figure 1, below) depicts a wide range of evaluative processes, from informal, experience-based after action review, or reflection, at one end, to formal evaluation with systematic application of the standards of the evaluation discipline at the other.

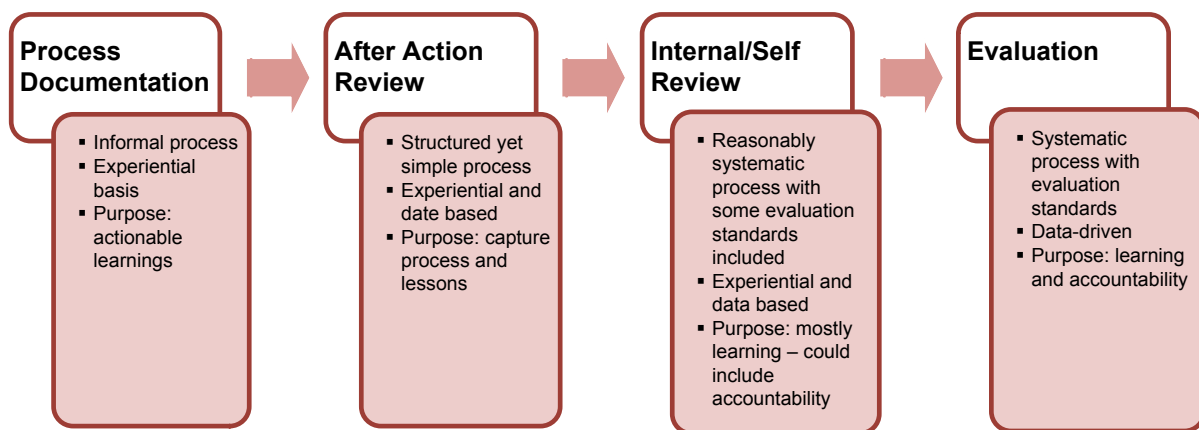


Figure 1: Illustrative Spectrum of Evaluative Processes (Church 2011b)

In situations like those described above, a **program quality assessment process** driven by a commitment to continuous learning might be a better fit than a formal evaluation. Professional standards for formal evaluations need not apply and the process would not have to be based on the OECD/DAC criteria (see OECD/DAC 2008, 2012). A program quality assessment would be an alternative to, but not a substitute for, a formal (summative or formative) evaluation. It would, however, require more external data than an after action review or reflection. We will discuss several processes that fall at various points on the spectrum above. Most of the Working Paper will expand on the idea of program quality assessments as applied to peacebuilding programming.

⁵ In addition to formal evaluations that use the program evaluation standards as the basis for determining quality, the evaluation discipline encourages evaluative thinking outside the evaluation process. Evaluative thinking is an ongoing process of “questioning, reflecting, learning and modifying” (Church 2011a, 7). In this process, programs ask, are we accomplishing what we wanted to accomplish? Why or why not?

III. PEACEBUILDING PROGRAM QUALITY ASSESSMENT

3.1 PROGRAM QUALITY ASSESSMENTS: KEY FEATURES AND INSIGHTS

As the use of quality assessments is still in its infancy in peacebuilding, the following considerations are best understood as initial ideas for a new tool for reviewing peacebuilding program quality, grounded in evaluative thinking. This proposed model could serve as the basis for pilot testing in the field.

A review of quality assessments in other fields like education, health, humanitarian aid, and community and social welfare programs reveals that all forms of program quality assessments have the following five characteristics in common:⁶

- The starting point is a review of the quality of the program’s design.
- They can be conducted either as self-assessments or external assessments.
- Criteria and standards for review are based on proven principles and/or “best practices” in the respective fields.
- Focus is on an ongoing process of evaluative thinking, learning and improving performance and effectiveness, rather than on empirical rigor and validity and accountability to an external authority.
- Capacity building is an integral element. Those implementing quality assessments are trained on how best to apply the assessments’ results and findings.

On the “Illustrative Spectrum of Evaluative Processes” in Figure 1 above (Church 2011b), a peacebuilding program quality assessment would sit near to internal/self review and formative evaluation, with some significant differences:

- The purpose of the program quality assessment would be primarily learning, not accountability.
- The program quality assessment would refer to criteria, standards or models for assessing program quality, in contrast to evaluation, which assesses whether a program is implemented as intended (process) and/or whether the program achieved its goals and had impacts on participants and on Peace Writ Large (outcome).
- The evidence standards for program evaluation would not apply; lower standards of rigor and validity in the data collection process would be possible.
- Capacity-building would be an integral part of the process, whereas in evaluations, the inclusion of capacity-building depends on the approach used.

⁶ See, e.g., <http://etools.highscope.org/pdf/YouthPQA.pdf>, http://www.childtrends.org/files/child_trends-2008_02_19_eval8programquality.pdf, and <http://forumfyi.org/content/measuring-youth-program-quality-guide-assessment-tools-2nd-edition>, all accessed last on December 5, 2012.

Table 1: Comparison of Program Quality Assessments with Other Evaluative Processes

	Reflection Exercise	Peacebuilding Program Quality Assessment	Evaluation
Purpose	<ul style="list-style-type: none"> Improve skills in program design Maximize potential for program to contribute to Peace Writ Large 	<ul style="list-style-type: none"> Learning about quality of design Adapt/improve program design and implementation 	<ul style="list-style-type: none"> Assess results/outcomes of program Learn about quality and value of program Accountability
When	<ul style="list-style-type: none"> During program/design phase At key moments (e.g., new phase) 	<ul style="list-style-type: none"> Mid-term Possibly at beginning or at end of project/program 	<ul style="list-style-type: none"> At end of project or program (summative) or Mid-term (formative)
Standards for assessment	<ul style="list-style-type: none"> RPP and Do No Harm concepts and tools related to impacts on Peace Writ Large and conflict sensitivity 	<ul style="list-style-type: none"> Best practices and research in program area (e.g., dialogue, trauma, media, etc.) and in strategy and program design (e.g., RPP findings) Program’s benchmarks 	<ul style="list-style-type: none"> OECD DAC evaluation criteria⁷ Contextually-relevant standards of achievement set by program/project
Standard of evidence (data needs)	<ul style="list-style-type: none"> Relies on knowledge and experience of participants 	<ul style="list-style-type: none"> Relies on program’s monitoring data, document review and some interviews with program team, partners and beneficiaries 	<ul style="list-style-type: none"> Standards for data collection/methodology should be followed Triangulation of evidence sought
Who conducts	<ul style="list-style-type: none"> Program team, possibly partners, selected experts and facilitator 	<ul style="list-style-type: none"> Internal team (self-assessment) or external evaluator with substantive knowledge of program area and facilitation skills 	<ul style="list-style-type: none"> Usually external evaluator with evaluation expertise and credentials; self-evaluation using same skills, standards and techniques as other evaluations also possible
Capacity building	<ul style="list-style-type: none"> Often part of capacity-building process Knowledge and skill in applying RPP and DNH tools required and often built into process 	<ul style="list-style-type: none"> Capacity building is an integral element—those implementing findings of quality assessments are trained on how to apply the assessment’s findings 	<ul style="list-style-type: none"> Capacity building often not an element
Resources	<ul style="list-style-type: none"> Minimal: fees for time, venue and funds for partner and expert & facilitator participation (if external parties or facilities are used) 	<ul style="list-style-type: none"> Medium: resources for evaluator, data collection process, as well as time, venue for discussion of findings 	<ul style="list-style-type: none"> High: resources for evaluator/evaluation team to design and implement rigorous data collection and analysis process.

⁷ The OECD/DAC evaluation criteria are impact, relevance, sustainability, efficiency and effectiveness (see OECD/DAC 2012).

Table 1 outlines some of the main differences between the proposed peacebuilding program quality assessments and formal evaluations, on the one hand, and a more informal reflection exercise, at the other end of the spectrum. The reflection exercise is similar to an “after action review” cited in Figure 1, yet it is intended to be a process for reflecting on the quality of program or project design and can be conducted in the context of training workshops (CDA 2011).⁸

3.2 When Should a Program Quality Assessment be Conducted?

A peacebuilding program quality assessment could be used at three stages of a program:

- During program design;
- For a program mid-term review; and/or
- At the end of a program.

At the program design stage, a program quality assessment can serve as a framework for reviewing and improving specific elements of the program’s approach.

During program implementation, a quality assessment can guide reflection on the progress achieved to date for a mid-term review. It allows collaborative learning and discussion on the strengths and unique opportunities, as well as its challenges and potential weaknesses. At this stage, a program quality assessment would promote consideration of whether to change the initial design and/or what needs to be changed for the next phase of a program.

At the end of a program, a program quality assessment would promote deeper reflection on what went well in the program or what would have to be changed in the next phase of a program or if a new program should be initiated. It might also inform the development of other similar programs by the same or other agencies, if it were made available publically.

3.3 WHO SHOULD CONDUCT AN ASSESSMENT?

A quality assessment can be performed as a self-assessment or as an external assessment. It could be conducted by:

- A local project/program team, with or without participation from headquarters;
- An external facilitator or expert in the role of a “strategic or learning coach” (in the understanding of Developmental Evaluation) (see Patton 2011, 24); and/or
- A donor agency.

A local project/program team might use the program quality assessment as a form of self-assessment. Such a self-assessment could be an important option when a team expresses resistance against a formal evaluation or where financial resources for formal evaluations are

⁸ In these exercises, workshop participants apply the full range of tools and concepts developed through CDA’s Reflecting on Peace Practice (RPP) Program (see section below) to review participants’ program proposals, drawing on the insights and experiences of colleagues from disparate organizations through the lens of RPP.

insufficient. At times, the local team is joined by someone from a regional or global office, especially in the case of larger international agencies.

An external facilitator may support a staff reflection on the program's design and results as part of an internal mid-term review process and/or at the end of a program. The focus would be mainly on collaborative learning and working with the team, rather than on collecting additional program data. In this case, an experienced evaluator may not automatically be best suited to act as "strategic coach" or external facilitator, as the skills are related but somewhat different. The role of the "strategic coach" is to help the organization to contextualize their program, facilitate learning and reflection, and support necessary adaptation and/or capacity building.

A donor agency may conduct a program quality assessment at the end of a program if the necessary financial resources for a formal evaluation are not available or when the conflict context is not conducive to a formal evaluation.

Both the external facilitator and donor agency have to ensure that the timing, setting/location of the program quality assessment, the selection of stakeholders involved and/or all other aspects of the process do not unintentionally or inadvertently increase the conflict dynamics or tensions between team members and/or stakeholders. In other words, the evaluation process itself must be conflict sensitive.

3.4 WHAT DATA IS NEEDED?

In general, the data reviewed includes the program documents and the information gathered through the program's monitoring and planning systems. In order to follow up on specific points and questions on the program design and logic, further interviews may be needed with the program team, partners, beneficiaries, or members of the communities, sectors or institutions the program is trying to influence.

During the program design phase, a review of the outcomes and lessons learned of similar programs and a review of conflict analysis will be useful. Depending on the quality of the conflict analysis (see Section 4.2.1 and Table 2 below), further conflict data may be needed.

For a program mid-term review, the focus will be on review of any monitoring data gathered by the program or project itself, the initial and updated conflict analysis, theory of change/overall program logic, outputs and outcomes of the program. The information would need to be supplemented by interviews with the program team, with beneficiaries and some other key informants in the area in which the program has been working. Analyzing this data, the level of empirical rigor both of the methods and of triangulation would not be as high as that required in a formal evaluation. For example, one need not invest in formal surveys and randomize respondents.

At the end of a program, a program quality assessment should provide analysis and evidence regarding lessons learned of the program. The quality assessment will inform what would have to be changed in the follow-up phase or a new program, should one be initiated. It might also

inform the development of other similar programs by the same or other agencies, if it were made available publically. The quality of data required would be similar to the mid-term review.

IV. Using RPP Concepts in Peacebuilding Program Quality Assessments

The Reflecting on Peace Practice Program (RPP) is an experience-based learning process that involves agencies whose programs attempt to prevent or mitigate violent conflict. During its first phase, from 1999-2003, it analyzed experience at the individual program level across a broad range of agencies and contexts, with the goal of improving the effectiveness of international peacebuilding efforts. RPP engaged over two hundred agencies and many individuals who work on conflict around the world in a collaborative effort to learn how to improve effectiveness vis-à-vis “Peace Writ Large”, or the larger societal peace. The findings of three years of analysis and consultation are presented in *Confronting War: Critical Lessons for Peace Practitioners*, which reviews peace practice, assesses elements that have been successful (or not) and why, and presents lessons on how to improve effectiveness. The lessons comprise a set of tools and concepts that are useful for conceptualization and planning of peace interventions at all levels. They help to answer the questions:

- *What* should we work on? Which of the issues or conflict factors is a priority?
- *Whom* should we work with? Which actors/stakeholders are most important?
- *Why* should we work on that issue with those people? Is the rationale for our chosen approach solid?

These findings, tools and frameworks have been enhanced, refined and supplemented since 2003 during RPP’s Utilization Phase, through ongoing work with peace practitioners and policy makers to apply the findings, as well as additional collaborative research on the cumulative effects of peacebuilding.

4.1 USE OF RPP TOOLS IN PEACEBUILDING EVALUATION

During this phase of RPP’s work, practitioners have been concerned with how to measure the impacts of peacebuilding efforts. In this context, RPP is exploring how its findings concerning *peacebuilding* effectiveness and impacts can be integrated with good evaluation practice. Church (2011b) examined how the tools and concepts that emerged from *Confronting War* and the Utilization Phase that followed its publication have been used in peacebuilding evaluation. The RPP tools and findings include:⁹

- **A systemic approach to conflict analysis**, including identifying and prioritizing key driving factors of conflict (KDFs) and the dynamics and relationships among factors;

⁹ For fuller explanations of the tools and concepts, see Anderson, Mary B., and Laura Olson. *Confronting War: Critical Lessons for Peace Practitioners*. Cambridge: CDA Collaborative Learning Projects, and CDA Collaborative Learning Projects, *Participant Training Manual* (2012), both available at www.cdacollaborative.org, last accessed December 5, 2012.

- The “**RPP Matrix**,” a tool for analyzing program strategies and their potential to affect Peace Writ Large. It can be used either in its full form, or broken down into its component findings: constituencies with which program work, either “key people” or “more people” and the linkages between the two; and levels of change sought, either “individual/personal” or “socio-political” change, and the linkages between the two.
- **Theories of Change**, including defining of goals, how achievement of goals affects the driving factors of conflict or peace, and the program theories explaining how the activities will “add up” to the achievement of the goal.¹⁰
- “**Building Blocks for Peace**” (formerly Criteria of Effectiveness), which represent a set of intermediate goals in peacebuilding that are useful for stimulating discussion of the aims of specific programs.
- **Linkages** across levels of change and people addressed.

In her review of twenty evaluations that used RPP in some way, Church (2011b) stressed several points that suggested the utility of RPP for a program quality assessment process:

- The evaluations that incorporated RPP were almost exclusively driven by a learning agenda, with accountability receiving little or no attention.
- The primary application of RPP in the evaluations reviewed was to assess the design of the project or program within an evaluation process.
- The RPP tools were generally applied “true to form” (not modified for the evaluation context).
- The majority of evaluations that incorporated RPP also integrated a participatory process, in order to apply the tool(s) or validate the conclusions.
- RPP tools and findings were used as an implicit set of standards against which evaluation findings were compared, at least as a basis for the development of recommendations for program improvement.

While RPP has been useful in helping frame lines of inquiry and in assessing relevance, it has been most useful to date in assessing and improving program design. RPP tools and concepts thus could be useful as a set of standards and as a framework for assessing program quality, either in the context of an evaluation or outside of it. Insights from the latest thinking on developmental evaluation¹¹ (DE) should also be taken into account, as they open promising avenues for peacebuilding quality assessments that incorporate RPP tools and concepts. In contrast to traditional evaluations, DE is most suitable in contexts that are:

¹⁰ The concept of Theories of Change (TOC) was not developed by RPP and has been used widely and in a great variety of contexts and fields. However, RPP was early to apply theories of change to peacebuilding as a way to bridge the gap between conflict analysis and programming. RPP has promoted the use of theories of change and has developed explanatory and training materials for peace practitioners. They are therefore taken into account as RPP tools, although the concept was not the product of RPP collaborative learning.

¹¹ Other names for developmental evaluation found in the evaluation literature include real time evaluation, emergent evaluation, action evaluation, and adaptive evaluation.

- Highly volatile and constantly changing;
- Socially complex adaptive systems;
- Difficult to predict;
- Open to risk and failure; and
- Open for learning and development (Dozois *et al.* 2010, 18).

DE proposes a “learning framework” that sets the direction for ongoing *learning* and strategic project *development*. This learning framework aims to map key challenges and opportunities in identifying a) what a team needs to pay attention to as they go forward and b) what a team needs to learn (Dozois *et al.* 2010, 31).

Clearly, it is beyond the scope of this paper to discuss the differences between traditional approaches to evaluation and DE (see, e.g., Gamble 2008, 61) and all the features of and steps involved in DE (Gamble 2008; Patton 2011). Yet, it is worthwhile to highlight two striking similarities between DE and RPP thinking and findings as we move toward considering RPP’s role in supporting quality assessment of peacebuilding programs.

DE and RPP promote similar ideas regarding the nature of social change and the role of an evaluation or evaluator. For comparison, the main equivalent insights of RPP are highlighted in italics below.

- DE provides feedback and generates learning and development and, therefore, helps to untie the internal “knots” and blind spots of an organization. Here, managing meaning and offering “interpretative frameworks” are as important as managing information (Patton 2011, 12-13). The developmental evaluator as part of a team is a “learning or strategic coach” infusing evaluative thinking into an organization and its staff.

Throughout their history, CDA and its RPP Program have placed the facilitation of collaborative learning and reflection at the heart of their processes. Furthermore, the RPP Matrix serves as a tool for mapping program activities, goals/outcomes and longer-term impacts on Peace Writ Large, which makes the practitioner aware of his/her strategic blind spots and untested and unrealistic theories of change.

- DE is driven by a “hunger for learning” and a “culture of the possible” (Dozois *et al.* 2010, 51). DE strives for ongoing strategic learning and innovation to respond to complex processes of social change that one cannot control. “Developmental evaluation is designed to be congruent with and to nurture developmental, emergent, innovative, and transformative processes” (Patton 2011, 7). Patton refers to double-loop learning as a systemic analysis which goes beyond identifying the problem and its solution (single loop learning) and unpacks the underlying assumptions, practices and values in the system leading to the problem in the first place (Patton 2011, 11). The emphasis is on generating ideas and “developing innovations” (Patton 2011, 22). Possible unanticipated consequences are integral part of the discussion in DE.

Both the RPP Matrix and the concept of theories of change facilitate a self-critical discussion of the untested and often unrealistic assumptions about social change driving peacebuilding programs and practitioners. The Action Asia Practitioner’s Forum illustrated how the RPP Matrix was used to guide discussions on desired and real outreach and impact (Action Asia

2011, 12). RPP's processes of systemic conflict analysis also support double loop learning and permit consideration of unintended consequences. Furthermore, the Do No Harm approach developed by CDA, while not an RPP concept as such, aims at capturing the possible inadvertent conflict-escalating and – deescalating effects of a project (Anderson 1999).

4.2 WHAT MIGHT A PEACEBUILDING PROGRAM QUALITY ASSESSMENT LOOK LIKE?

Based on the above discussion, this section discusses the possible shape of a peacebuilding program quality assessment, incorporating RPP concepts and thinking as standards for assessment. As RPP tools are concerned with strategic questions in relation to peacebuilding (how will the program affect Peace Writ Large), the use of RPP tools should be supplemented by standards and tools developed for the particular specialized substantive area in which the program is working (such as, dialogue, trauma healing, education, justice reform, etc.).

Applying the DE idea of a “learning framework” to the context of program quality in peacebuilding, the following RPP-based standards and lines of inquiry regarding program quality are proposed. An illustrative framework is presented in Annex A.

4.2.1 Conflict Analysis

“Analysis is not optional; it is essential and obligatory for peace work.”¹²

Quality peacebuilding program design and implementation should be based on an up-to-date, sound conflict analysis. This step is aimed at ensuring the “relevance” of the program: whether it is working on the right issue with the right people at the right time using an appropriate methodology in the particular context.

RPP has found that analysis is needed in order to avoid costly mistakes, find the correct program focus (addressing the right issues and people), identify priorities and strategic points of intervention, and match agency skills and resources to the situation. The evidence is strong that the more practitioners know about the conflicts they are trying to address, the more likely they are to identify effective avenues for work, and the less likely they are to make mistakes.

RPP has found, however, that agencies do not regularly do such analysis, and when they do, it is often partial analysis, focused on where, in a given context, the things they do can be useful or on where their expertise or theories can be applied. In addition, analyses are often too comprehensive (i.e., identify too many factors without prioritization or consideration of what is *driving* the conflict), static and often not updated. As a result, peacebuilding programs often “miss the mark,” because they are disconnected from analysis or designed to affect factors that are of lesser importance to the *conflict* dynamic.

While RPP does not single out any better framework or methodology for conducting conflict analysis, it does identify qualities of good conflict analysis—analysis that provides a sound basis

¹² This statement was offered by an RPP participant, as quoted in Anderson and Olson 2003, 51.

for effective program design and that is more likely to lead to “relevant” programming.¹³ These are listed as standards in the first column of Table 2, below. The second column outlines several shortcomings or gaps that RPP has found to be common in peacebuilding programming.

Table 2: RPP Analysis

	RPP-based standards	Common gaps/weaknesses or cautions concerning analysis
1	The analysis identifies <i>key driving factors</i> and <i>key actors</i> for conflict/peace. ¹⁴	<ul style="list-style-type: none"> • Analysis is <i>too</i> comprehensive: too many factors with no priorities identified • Analyzes entire context, but does not focus on conflict determinants; everything is seen as relevant to peacebuilding • Factors/issues are identified, but not dynamics among them or which are more/less important (priorities) • Analysis is implicit, and as a result not shared among team
2	The analysis considers what needs to be stopped (and who will resist) and what forces promote peace in this context.	<ul style="list-style-type: none"> • Analysis focuses on positive factors that might be strengthened, but does not consider countervailing negative forces • Analysis does not analyze what factors <i>actually</i> (rather than potentially) connect people or promote peace in this context
3	The analysis is updated and tested regularly/periodically.	<ul style="list-style-type: none"> • Analysis is performed once at beginning of program, but not updated • Conclusions about drivers and dynamics of conflict are not utilized to strengthen program
4	Analysis of similar program efforts either in the past or ongoing and any lessons from their results is included.	<ul style="list-style-type: none"> • Analysis does not identify, and program team has no knowledge of what has been tried before, with what results • Programs repeat failed approaches • Programs/projects duplicate efforts of others without added value
5	The program strategy builds on the analysis: identifies peacebuilding needs or points of leverage for change.	<ul style="list-style-type: none"> • Program goals and design do not address factors identified • Analysis does not enable program designers to identify what to do to change conflict dynamic
6	The scope of the analysis is appropriate (not too broad or narrow); and mitigates bias towards agency’s expertise or general beliefs about conflict.	<ul style="list-style-type: none"> • Analysis is performed to justify favored program approach (methodology, focus, constituency) • Analysis is based on beliefs about how to bring about peace generally (and not contextualized) • Analysis has omitted or excluded significant perspectives • Analysis exists at one level, but does not consider other levels (e.g., at local level, missing wider dynamics; national/regional analysis without

¹³ “Relevance” refers to whether the program the program is working on the right issues, at the right time, with the right constituencies, whether it is working on factors that are driving conflict or peace. For a fuller discussion of the dimensions of relevance and evaluating relevance in peacebuilding and conflict prevention, please refer to the first paper in this CDA paper series, Working Papers on Program Review & Evaluation, titled Evaluating Relevance in Peacebuilding Programs, by Mark M. Rogers, available at www.cdacollaborative.org. Last accessed December 5, 2012.

¹⁴ Key driving factors are elements/dynamics without which the conflict would not exist or would be significantly different. Key actors are people or groups that can significantly influence the conflict dynamics.

		local particularities; national analysis without international/regional dimensions)
7	The analysis process has been conflict-sensitive—considered potential harm it might cause?	<ul style="list-style-type: none"> • Team composition exposes team members or interviewees to danger • Team composition, behavior perceived as biased • Analysis process deepens polarization • Causes of conflict are contested among key parties—analysis process is fraught

4.2.2 Program Goals

Program goals need to be set in terms of desired changes. Many programs set goals for personal change, including attitudes, perceptions, personal behaviors, skills, and relations among individuals—and change at this level is often necessary, though rarely sufficient. However, programs that formulate goals as desired changes at the socio-political level are more likely to have impacts on Peace Writ Large. The socio-political level includes political, social and other institutions, as well as collective norms, attitudes and behaviors and intergroup relations.

While the desired changes do not need to be observable at the national level, programs/projects should, at the level at which they are operating, affect the creation of institutions (formal or informal), result in locally-driven peace agendas and action, or lead to *collective* attitudes and behaviors that reflect improvements in the key driving factors of conflict or peace at that level (e.g., increased security or perceptions of security, improved group attitudes or relations, resistance to violence, etc.).¹⁵

Table 3: RPP-based standards for program goals

	RPP-Based Standards	Common gaps/ weaknesses or cautions in goals
1	The goal addresses, directly or indirectly, key drivers of conflict or peace.	<ul style="list-style-type: none"> • The goal is not appropriate for conflict context; other goals may be more appropriate • The goal(s) addresses symptoms or consequences of conflict but not drivers, or it addresses factors of secondary importance.
2	The goal is stated as a desired change.	<ul style="list-style-type: none"> • Intended or expected changes from the effort are not clear • Goals are stated as activities, outputs or tasks
3	The goal is specific and realistic for the time frame—neither too broad (a long-term vision) nor too narrow (at the activity level).	<ul style="list-style-type: none"> • Goals are vague, grandiose and “overclaim” • Goals are expressed at a vision level • Goals stated are really processes (rather than the outcome of processes) or a series of activities • Program team is unable to articulate clear benchmarks, indicators or other signs that would help them know if progress is being made

¹⁵ The RPP Building Blocks for Peace emerged as intermediate-level criteria that indicate that a program is having or may have impacts on Peace Writ Large. See Anderson, Mary B., and Laura Olson. *Confronting War: Critical Lessons for Peace Practitioners*. Cambridge: CDA Collaborative Learning Projects, and CDA Collaborative Learning Projects, *Participant Training Manual* (2012), both available at www.cdacollaborative.org, last accessed December 5, 2012.

4	The goal is stated as a desired change in the socio-political realm. If not, there is an explicit longer-term strategy for effecting socio-political level change, or the program makes linkages to the activities of other agencies in the socio-political realm.	<ul style="list-style-type: none"> • The program/project seeks change at the individual-personal level only (attitudes, skills, etc.), and unrealistically “hopes” or assumes that changes at the socio-political will come about • The program goal at the individual-personal level is appropriate, but linkages to other programs or strategies for follow-on work to move to the socio-political level do not exist • The program assumes (without context-based evidence or conflict analysis) that a lot of work at the micro (community) level will somehow “add up” to significant changes at higher levels (Peace Writ Large) • Changes desired are not sustainable, big enough in scale or fast enough in this context
5	<p>The changes contribute to one of the following building blocks for peace:</p> <ol style="list-style-type: none"> Political institutions that address key drivers of conflict are created or reformed. Locally driven peace initiative/s address (indirectly or directly) key drivers of conflict/peace. People increasingly resist violence and/or provocations to violence. People gain increased security and/or a sense of security. Inter-group relations improve significantly (e.g., group attitudes, public opinion, social norms, public behavior). 	<ul style="list-style-type: none"> • Goals aim at individual attitude, skill or behavioral change only. • Teams do not agree on the type of change they are pursuing. • Program goals represent meaningful change, but it is not clear how the change might be sustained.

4.2.3 Overall theory of change

The overall Theory of Change, or the assumptions about how achieving the desired changes will address the key drivers of conflict and contribute to Peace Writ Large, should be clear, plausible and appropriate for the context.

Table 4: RPP-based standards regarding overall, or broad, theory of change

	RPP-Based Standards	Common gaps/weaknesses or cautions concerning overall theory of change
1	The Overall Theory of Change is explicit, with clear and understandable conceptualization of pathway to change.	<ul style="list-style-type: none"> • The overall theory(ies) of change is implicit, unstated. • It is unclear how the program, if successful, will affect key drivers of conflict or peace. • Theory of change is based on false assumptions about how change comes about in this context.

2	The Overall Theory of Change provides a plausible explanation of how achieving the goal will affect key drivers of conflict or peace.	<ul style="list-style-type: none"> Programs seek changes that are reasonable in themselves, but will ultimately fail to achieve sustainable peace (e.g., passing a law will not affect conflict drivers because agreement on principles and enforcement mechanisms are not being worked on; improving relationships may not necessarily facilitate greater political accommodation, tolerance or cooperation).
3	The overall Theory of Change is grounded in an understanding of how change happens in this context?	<ul style="list-style-type: none"> Ways in which change processes are different from context to context are not examined when program ideas or approaches from one context are applied in another.

4.2.4 Program strategy and logic

Assessing contribution to “Peace Writ Large” is difficult, as most peacebuilding programs are discrete efforts aimed at affecting one (often small) piece of the puzzle, and no one project can do everything. RPP has found that certain elements of program strategy and logic make it more likely that peacebuilding projects/programs will have influence on Peace Writ Large.

- The **program logic or theory of change** should be articulated and **clear, plausible, well-thought out** (with assumptions surfaced and examined), without leaps or gaps. It should also be grounded in the particular context, and specific enough to be testable.
- For peacebuilding programs to be effective, they must **link change at the individual/personal level to change at the socio-political level**. The individual/personal level includes attitudes, perceptions, behaviors, skills and interpersonal relations. The socio-political level includes relations among social groups, public opinion, social norms, societal institutions, and deeper elements embedded in social and economic structures and culture. RPP found that programming that focuses on or achieves change at the individual/personal level, but is never connected to or translated into action or results at the socio-political level, has no discernible effect on peace. In many cases, it is also important to link change at the socio-political level back to individual/personal level change—especially if the changes are to be meaningful (and not *pro forma*) and sustainable.
- Effective programs also **link work with “more people”**—people at many levels of society and in many sectors—to **“key people,”** people or groups that have the power or influence to decide for or against progress towards peace. Work that influences “more people” or “key people” but does not connect or link to efforts to affect the other has limited impact.

Table 5: RPP-based standards for program strategy and logic

	RPP-Based Standards	Common gaps/weaknesses in program strategy and logic
1	The activities “add up” to the goal—there is an explicit, rational and plausible link between	<ul style="list-style-type: none"> Program theory/logic is unclear or not explicit Program logic is weak/illogical or based on false assumption about how change comes about There are gaps or leaps of logic in the pathway to the goal

	RPP-Based Standards	Common gaps/weaknesses in program strategy and logic
	components of the program (input, output, outcome, impact) that is valid in this context.	<ul style="list-style-type: none"> Assumptions about how one change will lead to another (how the program activities will “add up”) have not been explored or articulated Team members proceed on very different assumptions about program goals, objectives and change processes Programs fail to account for key requirements (e.g., willingness & availability of participants) Assumptions have not been challenged or thought through
2	The program makes linkages between activities/changes at the individual/personal level and the socio-political level.	<ul style="list-style-type: none"> Program activities and changes are exclusively at the individual/personal level (attitudes, skills, relationships), with no strategy to translate these changes to socio-political change (either through follow-up activities or programs, or linkages with other efforts) Linkages or effects from the individual/personal to the socio-political level based only on “hopes” or assumptions
3	The program makes linkages between “more people” and “key people”.	<ul style="list-style-type: none"> Program focuses on the “easy to reach” with no strategy for reaching beyond to affect the “hard to reach”; “hard to reach” people or constituencies ignored Program has not incorporated strategies for affecting “key people” (if working with “more people”) or “more people” (if working with “key people”) Program works both with “more people” and “key people” and <i>assumes</i> linkage that may not occur, i.e. are unrealistic in this context Program believes it is working with key people when it is not (e.g., assumes government officials are “key” when they have little influence on the conflict; assumes people key to implementation of the program or to the mission of the agency are “key” to conflict; assumes victims of conflict are key) Analysis defines entire group of people (e.g., youth 15-25) as key but does not examine whether it will reach those likely to perpetrate violence
4	The scale and level of the outputs are reasonable in relation to the intended impacts and the size of the issue in this context.	<ul style="list-style-type: none"> There is mismatch between scale of goal (e.g., tolerance or reconciliation) and scale and level of output (e.g., number of participants, communities, etc.) The program is not “big” enough—does not have enough scale to have meaningful influence—and there is no strategy (either within or beyond the program or in conjunction with other efforts) for achieving meaningful scale
5	The program design has accounted for factors that could impede success, including ways social and	<ul style="list-style-type: none"> Program has not accounted for how the social and political systems will push back against efforts for change Program can/does achieve meaningful influence, but it is not clear how changes can/will be sustained—how, why and by

	RPP-Based Standards	Common gaps/weaknesses in program strategy and logic
	political systems might resist the changes the program is seeking.	whom?
6	The effort relates to other initiatives in the same or related areas of work, in terms of necessary complementarities, linkages or duplication of effort.	<ul style="list-style-type: none"> Program duplicates other efforts unnecessarily or fails to identify and link to complementary efforts in the same domain
7	The program is conflict-sensitive—it considers potential unintended negative impacts it might cause.	<ul style="list-style-type: none"> Program has unintended negative effects Program design has not examined common causes of unintended negative effects (the potential negative impacts of choices about beneficiaries, contractors, suppliers, location, distribution of benefits, etc.)
8	There is an M&E plan or feedback mechanism will provide timely, accurate and useful information about progress toward desired changes and about assumptions underlying theory of change.	<ul style="list-style-type: none"> There is no process in place for monitoring and testing the program logic and ensuring the program is not creating unintended negative effects Indicators are unrealistic

V. Conclusion

This paper has proposed the use of peacebuilding program quality assessments as an approach that reviews issues of excellence or inferiority of peace programming and addresses organizational needs for feedback, assessment and learning, short of a formal evaluation. We have suggested that peacebuilding program quality assessments can be based on standards indicated by RPP findings, tools and concepts. While there is some experience using RPP tools in various evaluative processes, this application to program quality assessments needs field testing.

Whether based on RPP or other standards, we are confident that program quality assessments have potential for broadening the international exchange on lessons learned and best practices regarding good program design in the peacebuilding field, including the testing of common theories of change.

References

- Action Asia. 2011. *Examining RPP as Tool for Evaluation. The Action Asia Experience. Results from the ACTION Asia Practitioner's Forum*. Colombo: Action Asia.
- Anderson, Mary B. 1999. *Do No Harm: How Aid Can Support Peace—or War*. Boulder: Lynne Rienner Publishers.
- Anderson, Mary B., and Laura Olson. 2003. *Confronting War: Critical Lessons for Peace Practitioners*. Cambridge, MA: CDA Collaborative Learning Projects.
- Blum, Andrew. 2011. "Improving Peacebuilding Evaluation: A Whole-of-Field Approach." USIP Special Report 280. Washington, D.C.: United States Institute of Peace.
- CDA Collaborative Learning Projects. 2012. *Reflecting on Peace Practice Participant Training Manual*. Cambridge, MA: CDA Collaborative Learning Projects.
- 2011. *Reflecting on Peace Practice Program Reflection Exercise*. Cambridge, MA: CDA Collaborative Learning Projects. [web access]
- 2010. *Reflecting on Peace Practice. Advanced Training of Consultant & Advisers. Resource Manual*. Cambridge, MA: CDA Collaborative Learning Projects.
- Church, Cheyanne Scharbatke. 2011a. "Peacebuilding Evaluation: Not Yet All It Could Be", *Berghof Handbook for Conflict Transformation. Section II*. Berlin: Berghof Foundation. Available at <http://www.berghof-handbook.net/articles/section-ii-analysing-conflict-and-assessing-conflict-transformation> (accessed December 5, 2012)
- 2011b. "The Use of Reflecting on Peace Practice (RPP) in Peacebuilding Evaluation. Review & Recommendations. Final Report." Cambridge, MA: CDA Collaborative Learning Projects. <http://cdacollaborative.org/publications/reflecting-on-peace-practice/rpp-books%C2%A0and-major-reports/the-use-of-reflecting-on-peace-practice-in-peacebuilding-evaluation> (accessed December 5, 2012).
- 2008. "Peacebuilding Evaluation: From Infancy to Teenager." *New Routes*. Vol. 13, No. 3: 3-6.
- Church, Cheyanne and Mark Rogers. 2006. *Designing for Results: Integrating Monitoring and Evaluation in Conflict Transformation Programs*. Washington, D.C.: Search for Common Ground.
- Davies, Rick and Jess Dart. 2005. *The "Most Significant Change" (MSC) Technique*. <http://mande.co.uk/docs/MSCGuide.pdf>, last accessed December 5, 2012.

- Dittli, Roland and M. Akhvlediani. 2009. *Promoting Dialogue for Regional Reconciliation*. Bern: Swisspeace, Center for Peacebuilding.
- Dozois, Elizabeth *et al.* 2010. *DE 201: A Practitioner's Guide to Developmental Evaluation*. Victoria/BC: The J.W. McConnell Family Foundation and the International Institute for Child Rights and Developments.
- Fetterman, David M. 2001. *Foundations of Empowerment Evaluation*. Thousand Oaks, CA: Sage Publications.
- Gamble, Jamie. A. A. 2008. *A Developmental Evaluation Primer*. Victoria/BC: The J.W. McConnell Family Foundation.
- Kawano-Chiu, Melanie. 2011. "Starting on the Same Page: A Lessons Report from the Peacebuilding Evaluation Project." Washington, D.C.: Alliance for Peacebuilding.
- OECD DAC. 2012. *Evaluating Peacebuilding Activities in Settings of Conflict and Fragility: Improving Learning for Results*. Paris: OECD.
- OECD/DAC. 2008. *Guidance on Evaluating Conflict Prevention and Peacebuilding Activities. Working Draft for Application*. Paris: OECD.
- Paffenholz, Thania. 2011. "Peacebuilding Evaluation: Assessing the Relevance and Effectiveness of Peacebuilding Initiatives: Lessons Learned from Testing New Approaches and Methodologies." Paper presented for the Annual Convention of the International Studies Association (Montreal, March 16-19 2011).
- Patton, Michael Quinn. 2011. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guildford Publications.
- 2008. *Utilized-Focused Evaluation*. Thousand Oaks, CA: Sage.
- Reimann, Cordula. 2012. *Evaluability Assessments in Peacebuilding Programming* Cambridge, MA: CDA Collaborative Learning Projects.
- 2010. "Nonviolent Peaceforce Evaluation of NP's Project in Mindanao, Philippines." Bern: Swisspeace.
- Rogers, Mark M. 2012. "Evaluation Impact in Peacebuilding Programs." Working Paper on Program Review and Evaluation # 2. Cambridge, MA: CDA Collaborative Learning Projects.
- 2011. "Supplement Evaluation Guidance. Catalytic Effects Component. Evaluability Assessment." Paper prepared for PeaceNexus Foundation and Peacebuilding Fund.

- Rossi, Peter and Howard E. Freeman *et al.* 2004. *Evaluation: A Systemic Approach*. Thousand Oaks, CA: Sage.
- Trevisan, Michael S. 2007. "Evaluability Assessment from 1986 to 2006." *American Journal of Evaluation*. Vol. 28: 290-303.
- Westley, Francis *et al.* 2006. *Getting to Maybe: How the World is Changed*. Toronto: Random House.
- Wholey, Joseph S. *et al.* (eds.) 2010. *Handbook of Practical Program Evaluation* San Francisco: Jossey-Bass.
- Woodrow, Peter and Diana Chigas. 2011. "Connecting the Dots: Evaluating Whether and How Programmes Address Conflict Systems", in D. Körppen *et al.* Eds. *The Non-Linearity of Peace Processes – Theory and Practice of Systemic Conflict Transformation*. Opladen/Farmington Hills: Barbara Budrich Verlag. http://www.berghof-peacesupport.org/books/sct_book_2011_Woodrow.pdf (last accessed December 5, 2012).
- Woodrow, Peter *et al.* 2008. "The Berghof Foundation in Sri Lanka: Resource Network for Conflict Transformation: Lessons Learned Evaluation Report." Berlin: Berghof Foundation.
- Yarbrough, Donald B. *et al.* 2010. *The Program Evaluation Standards. A Guide for Evaluators and Evaluation Users*. Thousand Oaks: Sage.

ANNEX A: Tool for Peacebuilding Quality Assessment

The following tool is an illustrative framework for organizing a peacebuilding quality assessment

In the left column are quality criteria or standards for a program design, based on the key findings, concepts and tools of RPP. The framework provides a series of questions divided into several broad categories: conflict analysis, program goals/outcomes, program strategy and logic, and unintended impact.

The second column asks for a rating of the program design, ranging from 1 to 3:

- 1 Program design meets none of standards
- 2 Program design meets some of standards, but there are significant gaps and weaknesses in several categories
- 3 Program design meets standards, with no significant gaps, although there may be room for improvement

The third column asks for evidence on the basis of which the assessment team or assessor has arrived at the rating. It should reflect strengths and opportunities—that is, where a program may fulfill quality criteria but might be strengthened further, as well as gaps and challenges—those aspects of a program that do not meet the standards or criteria of effective peacebuilding programs.

The fourth column provides for recommendations for improving the quality of the program design. This can include what an agency can do to sustain or improve the program design, what it should keep in mind or monitor while implementing the program, and suggestions on things that can be done to improve the quality of the peacebuilding program. Here the question of capacity building for responsible staff might also arise.

This framework tool may best be understood as a framework for an RPP-inspired learning process for peacebuilding organizations. It is not to be understood as comprehensive and exhaustive, but as preliminary reflection and food for thought.

I. CONFLICT ANALYSIS

Standards/Questions based on RPP	Rating	Evidence	Recommendations/Ways to Improve
Does the analysis identify <i>key driving factors</i> and <i>key actors</i> ?			
Does the analysis consider what needs to be stopped and what forces promote peace?			
Is the analysis updated and tested?			
Has the analysis identified and examined past or ongoing similar efforts and any lessons from their results?			
Does the analysis identify peacebuilding needs or points of leverage?			
Is the scope of the analysis is appropriate (not too broad or narrow, mitigates bias towards agency's expertise or general beliefs about conflict)?			
Was the analysis process conflict-sensitive?			

II. PROGRAM GOALS

Standards/Questions based on RPP	Rating	Evidence	Recommendations/Ways to Improve
Does the goal address, directly or indirectly, key drivers of conflict or peace?			
Is the goal stated as a desired change?			
Is the goal specific and realistic for the time frame—neither too broad (long-term vision) nor too narrow (activities or outputs)?			
Is the goal stated as a desired change at the socio-political level, or is achieving it part of a longer-term strategy for effecting change at the socio-political level? Does the program make linkages to efforts of other agencies at the socio-political level?			
<p>Do the changes contribute to one or more of the following intermediate building blocks of peacebuilding effect:</p> <ul style="list-style-type: none"> a) Creation or reform of political institutions that address KDF? b) Locally-driven peace initiatives addressing KDF? c) Increasing resistance to violence or provocations to violence? d) Increased security or sense of security? e) Meaningful improvement in <i>inter-group</i> relations? 			

III. OVERALL THEORY OF CHANGE

Standards/Questions based on RPP	Rating	Evidence	Recommendations/Ways to Improve
Is the Overall Theory of Change explicit, with clear and understandable conceptualization of pathway to change?			
Does the Overall Theory of Change provide a plausible explanation of how achieving the program goal will affect key drivers of conflict or peace?			
Is the Overall Theory of Change grounded in an understanding of how change happens in the particular conflict and context in which the program is being implemented?			

IV. PROGRAM STRATEGY AND LOGIC

Standards/Questions based on RPP	Rating	Evidence	Recommendations/Ways to Improve
Do activities “add up” to the goal? Are there explicit, rational and plausible links between components of the program that are valid in the context?			
Does the program make linkages between activities and changes at the individual-personal level and at the socio-political level?			
Does the program make linkages between “more people” and “key people”?			
Are the scale and level of the outputs reasonable in relation to the intended impacts and the size of the issue in this context?			
Has the design accounted for factors that could impede success (including ways social and political systems might resist changes)?			
Does the effort relate to or link with other initiatives in the same or related areas?			
Is the program is conflict-sensitive – does it consider potential unintended negative impacts it might cause?			

Will the M&E plan or feedback mechanism provide timely, accurate and useful information about progress toward desired changes and about assumptions underlying theory of change?			
--	--	--	--