The Use of Reflecting on Peace Practice (RPP) in Peacebuilding Evaluation

Review & Recommendations
Final Report

May 2011

Cheyanne Scharbatke Church
Besa Consulting: Catalyzing Strategic Change
Table of Contents

1. Introduction....3
2. A few key evaluation concepts....3
3. What is RPP?....8
4. Overarching Conclusions....8
5. How RPP is used in evaluation....8
   5.1 The parts of RPP being used....8
   5.2 The way in which RPP is being used....9
   5.3 Process Notes....14
6. Reactions to RPP Inclusion....16
7. Where should RPP go from here on evaluation....17
   7.1 Setting the stage for effective evaluation….17
   7.2 The application of the OECD-DAC evaluation criteria….19

Appendixes
   Appendix A: Methodology....20
   Appendix B: Interviewee List....21
   Appendix C: Evaluations or Related Documents Reviewed....22
   Appendix D: Author Biography....24

Acronyms

COE Criteria of Effectiveness
KDF Key Driving Factors
RPP Reflecting on Peace Practice
TOR Terms of Reference
1. Introduction
This report looks at how the Reflecting on Peace Practice (RPP) tools and concepts are used in peacebuilding evaluation and puts forth a set of recommendations on how RPP could add greater value to evaluation of peacebuilding. The process sought to gain a better understanding of the ways in which RPP is being used in project evaluations, to identify any unique contributions to the evaluative process and to identify opportunities for RPP to contribute to peacebuilding evaluation.

Conducted from October 2010 – February 2011, the project was commissioned by RPP and conducted by a long-term consultant to RPP who brings a wealth of evaluation expertise to the process (author biography available in Appendix D). The Review methodology followed a fairly standard course: concept note development, document review, interviews, discussion based on the draft conclusions with RPP Co-Director, Diana Chigas; all culminating in a draft report for review by the RPP Co-directors. Twenty evaluations were reviewed (see Appendix C), contributed mostly by members of the RPP Learning Community and 15 interviews (see Appendix B) were conducted. The vast majority of the evaluations were of projects with a few assessing programs (multiple projects). No strategy evaluations, regional or organisational evaluations were submitted for inclusion.

The status of the evaluations submitted for this exercise ranged from highly confidential to open access; therefore effort has been made to remove identifying variables from all examples and quotes. For more information on the methodology, please see Appendix A.

The report starts by laying out a few key evaluation concepts as a basis against which the Review was conducted. After which it summarises how RPP is being used in evaluation, followed by the reactions generated from RPP inclusion in evaluations from the perspective of the evaluand and evaluator. It concludes with some thoughts on where the process can go from here.

2. A few key evaluation concepts
How one has or should use RPP concepts in evaluation is influenced by what one considers to be evaluation. Agreeing on the boundaries of evaluation becomes a difficult task due to the inherent familiarity people have with the generic concept – to evaluate – resulting in an enormous spectrum of activities commonly deemed to be evaluation. What is evaluation gets further confused by the use of academic research data collection techniques such as interviews or surveys, in evaluation which often leads people to erroneously equate research with evaluation.

Though it is often not known in the international aid-development community, there is a significant body of academic theory and practice that is the basis of the evaluation field as a

---

1 This of course raises the issue of the applicability of these conclusions to evaluations behind project focused evaluations. It was beyond the scope of this study to assess RPP’s application in this regard.
discipline. Evaluation, like many disciplines including peacebuilding, has some developed norms and standards and an equal number of outstanding questions or debated positions.

One area that the discipline has developed a reasonable amount of agreement – though not uniform acceptance in terms of specific wording - is the definition of the term. As Scriven, one of the significant theorists in the field writes, evaluation is the “systematic and objective determination of the worth or merit of an object” (project, program). Implicit in this understanding of evaluation is the role of valuing or judgement; which is one of the key distinguishing features between evaluation and research.

There are different ways in which one can apply the judgement or valuing criteria inherent in an evaluation process. One of the primary ways is to use general standards against which to compare achievements of the evaluand to be able to determine if the work met, exceeded or fell below reasonable expectations informed by the context. For instance, if a community established a conflict resolution mechanism as a result of a peacebuilding intervention and that mechanism was predominately used for interpersonal dispute resolution, but not for issues that were drivers of the conflict such as land claims, standards would be one way to draw a conclusion on the worth of the intervention.

Where the evaluation discipline splits into camps is in the interpretation of this definition and specifically the breadth of steps that should be included within an evaluation. One camp – as exemplified by Jacob’s Five Tiered Approach or the Weiss/Chen Theory Based Evaluation approach – views the evaluative process starting with a review of the program design and working through to the understanding of results at different levels. This camp argues that there are key learning benefits to the inclusion of the full spectrum of the project cycle.

Conversely, the opposing view subscribes to a more narrow focus on assessing results or in evaluation language - the valuation of the merit and worth of the project. Scriven’s Goal Free evaluation approach and in some applications the Connoisseurship Approach exemplifies this perspective. This camp argues that by looking beyond merit and worth to program analysis and design issues, the evaluator’s focus and then resources become diverted and generally spread too thin. The argument continues that too much investment in the review of the program design component and not enough in the assessment (valuing) of the consequences and quality occurs when an evaluator attempts to engage in the full process.

An evaluation approach is the philosophy and process that underpins an evaluation.
There are more distinguishing features between evaluation and research than just the use of an evaluation approach and the inclusion of valuing. There are numerous other ways that the two fields can be distinguished. Summarised in Table 1: Research vs. Evaluation below, evaluation and research differ wholly or partially in their impetus, purpose, process considerations, stakeholder reaction, standards of data collection, analysis, deliverables and timeframe.

For each of the factors listed below the differences between research and evaluation are generally not absolute. For instance, it is common for both evaluations and research to have a primary report as the main deliverable. However the structure, audience and emphasis of the content of that report generally differ between an evaluation and a research report.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Evaluation</th>
<th>Traditional Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impetus</td>
<td>Stakeholders e.g. donor, NGO</td>
<td>Academic/Researcher</td>
</tr>
<tr>
<td>Purpose</td>
<td>Accountability &amp; Learning</td>
<td>Contribute to knowledge base</td>
</tr>
<tr>
<td>Process Considerations</td>
<td>Grounded in everyday realities of context &amp; organisation</td>
<td>Grounded in the needs of the research agenda and research standards</td>
</tr>
<tr>
<td>Stakeholder Reaction</td>
<td>Full range from a degree of anxiety, fear, resistance to enthusiastic embrace</td>
<td>Range from disinterest to willingness to contribute</td>
</tr>
<tr>
<td>Standards of Data Collection</td>
<td>“Good-enough” standard</td>
<td>Rigour and validity of methodology is critical</td>
</tr>
<tr>
<td>Analysis</td>
<td>Requires valuation of the findings</td>
<td>Findings are presented as the conclusions</td>
</tr>
<tr>
<td>Deliverables</td>
<td>Variety of possible products, structure and audience</td>
<td>Books and journal articles dominate</td>
</tr>
<tr>
<td>Timeframe</td>
<td>Snap-shot in time; length often determined by external factors e.g. budget, time, deadlines</td>
<td>Time is more often based on research question</td>
</tr>
</tbody>
</table>

Table 1: Research vs. Evaluation

One of the critical streams of work within the body of knowledge that constitutes the evaluation field is the work done on developing a discipline-wide understanding of what constitutes quality in evaluation. Standards have been developed that apply to the evaluation process and product

---

3 The American Evaluation Association (AEA) has been a significant contributor to advancing quality in the evaluation discipline. Though the program evaluation standards are American created, an analysis of evaluation quality standards internationally (e.g. OECD, UN, IDRC, Swiss Evaluation Society, Canadian Evaluation Society) show that these underpin the majority of standards documentation. That said, one cannot say that these specific five are recognized internationally, but rather the spirit of each is acknowledged.
and are not dependent upon what is being evaluated. Called the Program Evaluation Standards, the standards “identify and define evaluation quality and guide evaluators and evaluation users in the pursuit of evaluation quality.” (Yarbough et al 2010) The standards are grouped into five attributes of quality, which are:

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

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

- 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.

- Evaluation Accountability: standards intended to ensure the responsible use of resources to produce value.

When utilised, these standards are intended to catalyze greater evaluation quality. This is important as it can increase the credibility of the results and may galvanise greater use of evaluation findings at the program, policy or organisational level. In simple terms, more use of findings implies a closing of the learning loop which is of course one of the core drivers behind evaluation. Conversely, poor quality evaluation leads to low confidence in the process which can diminish the credibility of results and often leads to less utilisation (be it for learning or accountability) at all levels. In applying these standards to the average publicly accessible peacebuilding evaluation, one finds very significant gaps in their application⁴. Though a full range of quality does exist in peacebuilding evaluations, the average peacebuilding evaluation would not qualify as a high quality evaluation process or product as per these standards⁵.

---

⁴ For more on this please see Scharbatke Church, Peacebuilding Evaluation: Not Yet All It Could Be, in the forthcoming Berghof Handbook for Conflict Transformation, 2011.
⁵ Further research would need to be done to ascertain why these gaps exist in peacebuilding evaluation.
In addition to formal evaluation processes which use the program evaluation standards as the basis of quality, the evaluation discipline also emphasizes other processes that promote active reflection. The evaluation field promotes the importance of evaluative thinking within and outside of the evaluation process. Evaluative thinking is a means of thinking or of viewing the world. It is an ongoing process of questioning, reflecting, learning and modifying: asking oneself “what are we learning and how can we use those lessons to improve our performance?"

As stated above, one can find evaluative thinking underpinning a wide spectrum of evaluative processes, not just formal evaluations. The Illustrative Spectrum of Evaluative Processes found in Figure 1 below locates formal evaluation with a systematic, rigorous application of the norms and standards of the evaluation discipline at one end of the spectrum and more informal, experiential based (rather than data based) exercises at the other end such as an After Action Review. It does not attempt to lay out all of the possible evaluative processes, but rather give a flavour of the rich variety of options available to social change practitioners seeking to apply evaluative thinking to the benefit of their work and ultimately their beneficiaries.

The utility of understanding the range of processes based on evaluative thinking is significant; not the least of which is providing a means of benefiting from evaluative thinking without always having to implement a full evaluation process. The determining factor to which reflective activity is used – After Action Review or an evaluation - should be based on what the organisation is seeking to achieve through the process.
3. What is RPP?
For the purposes of this Review, a determination of what ‘counts’ as RPP had to be made to identify where it appears or not in evaluations. As a result, the following were deemed to be representative of RPP:

- the matrix: in full form or broken out by concepts: key and more people; individual and socio-political level change,
- criteria of effectiveness (COE),
- key driving factors in conflict,
- determination of ‘adding-up’,
- linkages,
- systems approach to conflict analysis.

Theories of change was felt to not be RPP specific and thus did not qualify as an RPP concept though it is of course part of the RPP discourse.

4. Overarching Conclusions
- The evaluations that incorporated RPP were almost exclusively driven by a learning agenda; with accountability receiving little to no attention.

- The primary application of RPP in the evaluations reviewed was to assess the design of the project within an evaluation process; placing them squarely in the broader interpretation of what constitutes evaluation.

- RPP tools were generally applied ‘true to form’; not being adapted or modified to the evaluation context thus supporting somewhat the concerns of the evaluation community who feels that design review is outside of the auspice of evaluation due to the diversion of focus.

- The RPP tools themselves do not – at present - lend themselves to being evaluative as they are felt to be more descriptive than analytical. However they can help support evaluative thinking.

5. How RPP is used in evaluations
One of the main thrusts of this Review was to map how RPP was currently being utilised in peacebuilding evaluations. This section looks at which parts of RPP are being used, to what end they are being used and finally some process notes.

---

Please see the 2009 Reflecting on Peace Practice: Participant Training Manual for definitions and explanations of these concepts.
5.1 The parts of RPP that are being used

Those conducting evaluations – either as commissioners or evaluators - approach RPP as a menu of tools and concepts from which they can pick to suit their purposes. The following list indicates those tools/concepts used; represented from most (#1) to least (#6):

1. The matrix either in full or broken out by component (e.g. key and more people)
2. Key driving factors
3. Criteria of effectiveness (COE)
4. Linkages
5. Determining ‘adding-up’
6. Systems approach to conflict analysis

Across the board, the application of the selected tool stays true to the way in which it is laid out by RPP, with no one in this review attempting to adapt or modify the tools to meet their purpose or the evaluation process.

5.2 The way in which RPP is being used

RPP tools and concepts are being used in three primary ways – listed from most to least; as a de-facto standard, as a frame for inquiry and as a means to assess relevance. There are also a few unique applications that are summarised at the end of this section. After each use of a RPP tool is described, a comparison is made with ‘the norm’ on the issue within evaluation discipline in an attempt to highlight similarities and differences.

5.2.1 As a de-facto standard

The most common way in which RPP tools are used is as a standard against which evaluation findings are compared in order to draw evaluative conclusions about effectiveness. For instance, a program strategy is deemed ‘good’ if it is responding to key drivers of the conflict (the COE pre-requisite); while it is seen in need of improvement if not addressing key driving factors. Here the ‘necessity to address key driving factors of the conflict’ serves as the standard or basis of the judgement.\(^7\)

The RPP concepts that were most commonly used as ‘standards’ were:

1. Who the program was addressing – more vs key people (matrix)
2. What level of change – individual vs socio-political (matrix)
3. Existence of explicit linkages with other organisations

---

\(^7\) There are no internationally accepted standards for effectiveness in peacebuilding as such where evaluative conclusions are attempted in the average peacebuilding evaluation (which is not the norm) they are generally based on implicit norms or expectations possessed by the evaluation team.
4. Key driving factors of conflict

The following are a few illustrative examples of the evaluative conclusions that have been drawn based on RPP standards.

- **Linkage as the standard**: “Finding 10: The project is not well enough connected to other like-minded organisations.”
- **COE as the standard**: Finding 12: the project addresses directly one of the criteria, has the potential to address two more with minor changes, and has also a clear potential of including the remaining two criteria if a new line of activity is added. In sum, this shows that the project focus is well situated and has a very high potential to contribute to effective peacebuilding.”
- **The matrix as a standard**: “The movement and expansion of [X program] within the matrix, together with the reinforcing effect of other programmes of support, strongly suggests that the likelihood of meaningful impact has increased.”
- **Key driving factors as a standard**: “The consensus building process worked in a way which did not however always facilitate the identification and prioritization of key driving factors of the X conflict.”

An **evaluative conclusion** embodies the judgment about the finding, compared to a research finding which simply states the sum of the data on the particular issue.

In using the RPP tools as standards against which to assess a program design for effectiveness some evaluations are clear that this is an assessment of potential or the likelihood of a program contributing to peace writ large. For instance:

- “If the X program is addressing some of the key driving factors of conflict then its structure and strategy give it the possibility of real impact and influence” (emphasis added)

Others use the tool or lesson as evidence of effectiveness; more of a statement of proof that the program did affect change. Examples include:

- “Finding 9: The projects activities are of limited effectiveness as they do not try to induce change on both personal and socio-political levels.”
- “It shows us how we can measure our effectiveness and helps us do it.”

What is rarely clear from the evaluation reports is whether additional data was then collected to confirm this conclusion. This could be due to insufficient detail provided in the methodology or lack of inclusion of evidence to support statements, even though the data was collected.

None of the evaluation reports acknowledge the use of RPP as the standard that will be used in explicit evaluation language; for instance the term standard is never used. Where it came up in the interviews amongst those who had used RPP and had somewhat of an evaluation background the application of RPP as a set of standards was reacted to with a degree of discomfort. Here are a few of the pertinent statements from the interview process:
“I am uncomfortable with the idea that an evaluator comes and he has a set of benchmarks in his backpack and he is judging us by these standards and I don’t know if RPP can be called accepted standards for evaluation.”

“I have a degree of discomfort with this. I feel they haven’t been tested broadly enough. People who have put them in question simply ignore them and find something else and people who it resonates with use them.”

“...use of them as standards in evaluation – are the frameworks evidence tested – you can’t say I am evaluating this against RPP criteria which are in fact guesses – hypothesis derived against a biased sample.”

On the flip side of using RPP as standards is the use of RPP concepts as the basis of recommendations. In other words if a program was not deemed to meet a standard (RPP lesson) then the recommendation would generally be something to correct the design so that the standard would be met. For instance, if the program design was seen to only work in the individual realm then one of the recommendations would be to ensure that the new thinking included movement into the socio-political realm.

Standards in the evaluation world: Comparing findings against a standard to conclude an evaluative statement is a classic evaluative function as it is the process of assessing or valuing. Different sectors have more or less standards in existence, for instance the humanitarian field has given significant attention to standard creation over the past several years.

5.2.2 As a frame of inquiry for the evaluation

RPP lessons can be found in the lines of inquiry (evaluation questions) which form a guide or direction to what areas the evaluation explores. As one interviewee stated, “These [RPP concepts] guided their inquiry. It told us what to be curious about.” In this way RPP is informing what to evaluate but not how one does the evaluation.

In some cases this framing can be identified in the evaluation Terms of Reference (TOR) themselves. In one case a donor has included numerous RPP concepts in their guidance note for those developing Terms of References. The Guidelines on evaluating the success of projects/programmes in evaluation priority sector “Peace Building and Crisis Prevention” has incorporated RPP through suggested lines of inquiry including:

Under the criteria of Relevance and Appropriateness it asks “To what extent are driving factors and key actors of the conflict addressed?” [p4]

---

8 It is impossible to identify the extent of this use of RPP as so few evaluations appended their TORs to the evaluation report.
Under the criteria of Impact it asks:

- “Has the development measure changed the attitudes and conduct of target groups and relationships between the conflict parties?” This is criteria of effectiveness (COE) # 5
- “Has the project or program promoted local peace initiatives and generated its own momentum for peace?” This is COE # 2.
- “Have state and civil society institutions been advised and strengthened in order to implement measures to prevent violence, promote dialogue, conduct reconciliation processes etc.?” This is COE # 1. [All quotes from p 7.]

Of note, one sees far more use of the Criteria of Effectiveness in a framing context than in use as standards. However, interviews highlighted more challenges with the use of COE than with the other lessons or tools. On first pass practitioners and evaluators tend to like the questions that the COE can be translated into, however evaluators found responding to these questions raised several issues. As one interviewee noted, “Even if it [the COE] generates questions which look good they are a bit difficult from an evaluation perspective.”

First, there is a tendency of the team being evaluated (the evaluand) to cherry-pick those COE that they feel are pertinent to the project at hand and only include them in the TOR. Only assessing a limited number of COE ignores the cumulative – more is better – nature of the COE. For instance, one interviewee explained how they applied the COE, “We use the COE to measure the effectiveness of different projects. We found that not all 5 COE fit our projects – whatever criteria fit in our projects we include it.”

Second, those conducting the evaluation struggled to determine what constitutes achievement of any one of the COE. A general sense existed that there needed to be more direction on what was necessary in order to determine that the COE had been successfully achieved. In other words what degree and scale of change is necessary within each of the COE for the evaluator to conclude this COE has been met. However the sentiment was also expressed that each COE needs to be contextualised in order to have meaning for an evaluation.

Third, it was mentioned by a very few that assessing change against the COE is very difficult given the absence of baseline information. Finally, there is a sense that the 5 as they currently stand are not ‘even’ or of an equivalent significance.  

---

9 Though these challenges were identified with the application of the COE, to date no one has developed a means of addressing or responding to them in their evaluations.
Frames of inquiry in the evaluation world: Lines of inquiry/evaluation questions are a standard way in which to frame an evaluation. Some sectors would also have sector specific evaluation questions.

5.2.3 As a means to assess relevance

RPP tools, notably the importance of programming addressing Key Driving Factors (KDF) and the need to do a conflict analysis in order to ensure programming does address KDF, are being used in the determination of relevance. This application of RPP was found the least in the evaluations reviewed and therefore has the least material to summarise. Further most references of this application appear in the evaluation reports as statements with limited explanation or description to support the statement. For instance, none of the evaluations which used these RPP methods in determining relevance provided any form of criteria or articulated methodology to explain why the identified key driving factors were deemed ‘valid’.

In terms of methodology to conduct or verify the conflict analysis, only the CDA led evaluations used systems-based conflict analysis, while the non-CDA evaluations appeared to use interviews and staff small group discussions in order to verify the analysis. Most of the evaluations then indicated if the programming responded to these key driving factors or not, however none attempted to assess whether the strategy/activities selected were the ‘best’ response possible.

Assessing relevance in the evaluation world: The inclusion of an analysis as the basis for a relevance assessment is standard practice in the evaluation discipline. The importance of the context is often stressed in evaluation literature and many evaluation approaches include it as an overt step, such as the Stufflebeam CIPPS approach. It may be that the type of analysis – conflict analysis – is different, however I am not familiar enough with the evaluation of domestic social change programming such as education or domestic violence programming to know if they too have a tailored analytical approach.

In terms of the use of a systems-based analysis, this can also be found in the evaluation world, however it would not be considered a ‘norm’ of the practice. Systems as an approach in evaluation started to gain momentum approximately five years ago and the application of systems is still mostly seen within those who do systems-based evaluations with two exceptions. First is the Outcome Mapping approach, pioneered by the IDRC, which is based on systems thinking and complexity theory. This approach, though not new, is gaining increasing attention in the international community. Also in the past year there have been notable references to
systems in the broader evaluation literature, such as Quinn Patton’s new Developmental Evaluation.

5.2.4 Unique applications of RPP concepts
Three people used RPP concepts in a way not seen in any of the other efforts reviewed in this review. Given that both overarching lessons and innovative ideas are of utility at this point in the process they are noted here.

1. The COE were used as a framework to describe impact for the end of year reporting of one organisation. This is more of an example of RPP use for monitoring than for evaluation, but if there was an evaluation of this organisation, pending the quality of data, the data gathered could make a valuable contribution to the evaluation exercise.

2. One evaluator utilised the key people – more people distinction as a principle in their ‘sampling plan’ for the evaluation methodology. Not surprisingly this experience showed that gaining access to more people was far easier than gaining access to key people in the evaluation. That said, the information gathered was extremely useful for the evaluation with each grouping offering differing perspectives.

3. One evaluator explicitly reframed an evaluation intended to be a formative (midterm) in its TOR into a “joint learning process” that was explicitly geared “towards improving the quality of a project based on joint learning and open reflection by all involved stakeholders.” This meant the focus was on the quality of the design and the intended or potential effectiveness. In this situation the project under review had completed one year of a three year grant and there were some concerns about its progress and a reflective effort was deemed to be more useful. Though the focus was different the steps taken – document review, small group discussions with staff, interviews, workshop – did not differ drastically from the average process found in the evaluations within this review.

5.3 Process notes
There are two different interpretations of process in this section. First is the way in which RPP gets included into the evaluation and the second refers to the process used by the evaluators in conducting the evaluation.

There is no one way that RPP gets incorporated into an evaluation. As has been described earlier, in some cases RPP tools and concepts are directly mentioned in the evaluation TOR. In...
other situations, RPP appears to be injected by the evaluator in the development of the evaluation methodology; as can be seen in a few cases in RPP inclusion in Inception Reports. Finally, there were also a number of instances where the inclusion of RPP ideas was more emergent; they were included as the evaluation unfolded. As data was collected or analyzed, it became apparent to the evaluators that a RPP tool or concept would be beneficial to the process and therefore was incorporated.

In terms of the second interpretation of process, the way evaluators conducted the evaluation, there were some commonalities. The process commonly started with a document review, followed by data collection which was almost exclusively interviews with key informants or small group interviews, followed by analysis and the development of a draft report\textsuperscript{12}. None of the evaluations described an analytical method or utilised rubrics to focus the analysis.

At different points in that process, the majority of evaluations incorporated a participatory element, generally with the implementing team, in order to apply the RPP tool or validate the conclusions of the tool.

This participatory element took on various forms including:

- formal 3 day workshop style trainings on the RPP tools and concepts to ‘teach’ the concepts,
- short (1/2 – 1 day) overviews of the RPP core lessons to provide a conceptual grounding,
- small group discussions to validate the analysis done with the RPP tool and
- group meetings to assess how the work fit into RPP tools.

In one case the tools were taken beyond the staff teams and incorporated into interviews with community members – so not in a participatory manner.

There was a sense amongst some of the participants in this review that one had to include an additional training on RPP in the evaluation process in order for the application of the RPP tools to be a useful element. A few raised concerns about this perceived prerequisite in two forms. First, because it adds another piece to a process that is generally already underfunded and working under tight time frames. Secondly it was questioned as to whether this is an indication of a weakness in the tools themselves that they ‘require’ a training in the application before they are adopted\textsuperscript{13}.

\textsuperscript{12} This is a fairly classic ‘research’ style process. One reason for the consistency in processes could be a perception that evaluation and research are the same activity. It could also be a symptom of lack of time, underfunding or expectations of the client that this process is ‘the way’ to do evaluation.

\textsuperscript{13} Whether this is a weakness in the tool or in the way in which the tool was described in the report is hard to know given the information at hand.
In terms of data collection methodology no baselines or comparison or control groups were utilised in any of the evaluations. There was almost no use of quantitative data collection or mixed methods. Further there was limited explicit reference (though maybe it just wasn’t documented) to the use of triangulation either through sources or data types. Finally, methodological limitations were rarely explained in evaluation reports.

**Process in the evaluation world:** Evaluation process is one area where there are many differences between the evaluations included in the Review and the evaluation world. An evaluation approach underpins the process implemented in an evaluation – one part of which is the data collection – but it also can include stakeholder engagement, capacity building, standards creation, development of analytical frameworks amongst other steps. In the evaluation world there are three camps of approaches – those that are focused on utility, those that emphasis methods and those that have values playing a central role. Each camp has numerous different approaches within it.

The evaluations covered in this review made no reference to the evaluation approach underpinning the process. One can infer from the tenor of many reports and decidedly from conversations with those conducting the evaluations that there is a desire for the evaluation to be of utility to the group being evaluated (the evaluand). If this interpretation is accurate, there are numerous evaluation approaches that prioritise utility of the evaluation process and product, incorporate extensive participatory methods and/or build capacity throughout the evaluation that do not appear to be used to inform these evaluations.

In terms of data collection, the program evaluation standards (described on page 6) provide guidelines to aid in the interpretation of quality. When using the data collection pertinent standards, gaps can be identified between them and the evaluations used in this review. For instance, mixed methods – where feasible – are perceived to be the best data collection option. Or including a description of the methodological limitations in the evaluation report is seen to be a required component to enable all readers of the report to understand the strengths and weaknesses of the evaluation design.

**6. Reactions to RPP inclusion**

It appears that the evaluand’s reaction to the inclusion of RPP was directly related to the evaluand’s previous exposure to RPP. Generally, participants in the study noted quite positive reactions from those evaluated if they were familiar with RPP – either prior to the evaluation or through participatory elements in the evaluation. If they were not familiar with RPP or the participatory element was not ‘adequate’ then there was some push-back to the use of RPP, predominately when they were the basis of conclusions (application of standards).

The COE generated some specific reactions from those interviewed:
“It sounds when you talk to the people about this [COE], it sounds like it is putting it all in a box. They elicit more questions than answers for me. For people who have been not so much involved they are not so useful.”

“They [evaluation commissioners] were very keen on the COE – when we used them they became less keen. They saw that those criteria disadvantaged their work a bit. Some of the things they hoped to say didn’t seem so likely when those were used.”

“They [the client] were not keen. They acknowledged them a bit but did not welcome it. They saw other COE that they were interested in.”

Those in the role of evaluators noted that they liked the inclusion of RPP as they felt they were offering something tangible, useful or new to the group being evaluated; a sense of clear value-add from the process was expressed. Some of the evaluators also noted that RPP allowed them to identify what was going well and then what could be improved. This recognition of achievement was felt to be very important to not discourage those implementing the work on the ground. As one interviewee reflected, “This is a way to acknowledge the good work and then place it in a larger context of what the impact might be.”

7. Where should RPP go from here on evaluation?
There are two strands of work that RPP could explore that build from the existing application of RPP in evaluation and reflect gaps in the broader peacebuilding evaluation field. First, RPP can build on its existing work to support and strengthen the quality of peacebuilding intervention design in a way that sets the stage for more effective evaluation. Second, RPP could contribute to the understanding of how to apply some of the OECD-DAC evaluation criteria.

7.1 Setting the stage for effective evaluation
RPP has experienced significant resonance with practitioners on the issue of program design. As the 2009 RPP Participant Training Manual states “The lessons comprise a set of tools and concepts that are most useful for the conceptualization and planning of peace interventions at all levels.” Therefore one avenue would be to play to the strengths of RPP and locate their application as a precursor to evaluation. This could be done in two ways: increasing evaluability and quality reviews.

7.1.1 Using RPP to make projects more evaluable:
One of the common challenges in peacebuilding evaluation is the project under review is not ‘evaluable’. This leaves the evaluator with one of two choices – spend limited resources on restructuring the project to make it evaluable (but in so doing reduce the...
resources available to the actual evaluation) or move forward and simply draw the conclusion in the report that they were unable to assess the project due to the weakness in design. Neither option is optimal; generally resulting in a sense of frustration with the resulting evaluation.

There is an opportunity to take action before the evaluation is started to determine if the project is evaluable. This step, called an evaluability assessment, is reasonably well known in the evaluation world, but has not had uptake in the peacebuilding community. In essence this step assesses if an evaluation should go forward. Criteria commonly utilised in this assessment include:

- Low dosage – program is weak
- Immature – program is continuously evolving
- Amorphous – no explicit or credible logic or theory
- The good cause – no discernible goals or objectives
- Known quantity – the impact is already well known
- Poor delivery model
- Unethical
- Nothing to compare to
- A negative finding cannot be accepted

Where these criteria showed that peacebuilding programs were not evaluable (for instance, they were deemed to be low dosage, immature, amorphous or the good cause) they would be excellent candidates to go through a RPP process. Through this application of RPP, the program would become evaluable and be in a far stronger place to experience a useful evaluation process.

It is suggested that a process be developed that blended RPP with evaluability assessment for peacebuilding programs; creating an RPP infused evaluability assessment. This process could be a standalone application or the first stage in an evaluation. Either way it should not be confused with the evaluation and must be resourced adequately, both in terms of budget and appropriately skilled implementers.

7.1.2 Using RPP to develop a peacebuilding project quality assessment process:

Quality is something that is often highly valued and sought after, but rarely specifically defined in the peacebuilding world. In theory, high quality should bring greater effectiveness which is the change in the world that we seek. Though evaluations are well placed to assess quality, there is also a need for a quality assessment process that sits

---

14 Simply because someone is a good evaluator would not necessarily make them skilled at a RPP infused evaluability assessment. Likewise, the person best placed to do a RPP infused evaluability assessment may not make a good evaluator.
outside of a structured and systematic evaluation process. In instances where there are
significant concerns regarding the effectiveness of peacebuilding programming or high
resistance to evaluation, a quality assessment process that is driven by a commitment to
continuous learning and capacity development might be a better fit. A process could be
developed that is grounded in evaluative thinking while building upon RPP tools and
concepts.

Referencing the evaluative thinking spectrum (see page 7) a quality assessment process
would sit on the spectrum very near to formative evaluation, but would be distinguished from
it in a number of ways. First the capacity building element would be a standard component
in a quality review whereas its inclusion is based on the evaluation approach adopted in a
formative evaluation. Second, the program evaluation standards would not apply. Third,
there would be instances where the process did not collect any additional data but simply
worked with the team. This would minimise the expectations of rigour and validity in the
data collection process. Finally the OECD DAC evaluation criteria would not be the basis of
the process, but rather a quality review of design.

8. The application of the OECD-DAC evaluation criteria
Far more work is necessary to understand the best way(s) to apply the evaluation criteria outlined
has a particular ‘value-add’ is in relation to determination of Relevance and Impact.

- **How to determine Relevance in peacebuilding evaluations:** An area that RPP is beginning
to contribute is the relevance assessment in peacebuilding evaluations. However there is
much yet to be developed; particularly how an evaluator determines the quality of the
conflict analysis and appropriateness of the resulting strategy. This is the evaluation side of
the bridge between analysis and strategy work in the design stage. This is an important gap
in the current practice and one of the few areas that peacebuilding evaluation may indeed be
unique from evaluating other interventions.

- **How to determine Impact in peacebuilding evaluations:** The Criteria of Effectiveness are
best placed to contribute to the Impact assessment within an evaluation. However as has
been highlighted already evaluators find the practical application of the COE particularly
difficult. Making the COE ‘evaluation applicable’ could be a significant contribution to
peacebuilding evaluation.
Appendix A: Review Methodology

This Review started with a call for evaluations from the RPP Learning Community; supplemented by the CDA led evaluations or evaluations CDA had on file conducted by others. This gathering of evaluations resulted in 20 evaluations or reviews being submitted plus 23 other pertinent documents. Of the 20 evaluations 7 were conducted by one of the RPP Co-Directors. Consultants who have been closely identified with RPP and where the Co-Director was an advisor only on the process were still counted as non-CDA led evaluations. Three additional documents; an article, a set of guidelines and notes from an in-house review of RPP application were also contributed. These documents were assessed for their application of RPP.

All those who submitted an evaluation to the process were asked to participate in a telephone interview. This list was supplemented to expand the coverage with a few evaluators who are known to be familiar with RPP. As a result 15 interviews were conducted as part of this process.

Interview break-down based on perceived nationality:
- 2 Southeast Asia
- 6 North Americans
- 4 Europeans
- 2 Africa
- 1 Latin America

Self-assessed professional description, recognising that many people offered up hyphenated responses e.g. practitioner-consultant:
- 8 practitioners
- 3 consultants
- 2 evaluators
- 2 researcher/academics

The interviews were followed by pre-draft discussion with Diana Chigas where tentative findings were explained and discussed. The feedback from this conversation was blended into the subsequent analysis and draft report writing. Two further meetings with the co-directors took place to discuss the findings and the recommendations for going forward.

---

15 Of the 13 who do not describe themselves as evaluators, the majority do not have training in evaluation based on the evaluation discipline. This had an impact on the inquiry as we were not able to explore perceptions of benefits, drawbacks or differences from evaluation as practiced in the evaluation discipline.
### Appendix B: Interview List

<table>
<thead>
<tr>
<th>Name</th>
<th>Title</th>
<th>Organisation</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emery Brusset</td>
<td>Director</td>
<td>Channel Research</td>
<td>Switzerland</td>
</tr>
<tr>
<td>Christian Bugnion</td>
<td>Director</td>
<td>Subur Consulting SL</td>
<td>Spain</td>
</tr>
<tr>
<td>Bridget Butt</td>
<td>Programme Director</td>
<td>Change Agents for Peace International</td>
<td>Kenya</td>
</tr>
<tr>
<td>Susanna Campbell</td>
<td>Research Fellow</td>
<td>Centre on Conflict, Development, and Peacebuilding (CCDP), The Graduate Institute PhD Candidate, The Fletcher School, Tufts University</td>
<td>USA</td>
</tr>
<tr>
<td>Shiva Dhungana</td>
<td>Design, Monitoring &amp; Evaluation Specialist</td>
<td>Search for Common Ground</td>
<td>Nepal</td>
</tr>
<tr>
<td>Roland Dittli</td>
<td>Head of Peacebuilding Analysis &amp; Impact Programme</td>
<td>Swisspeace</td>
<td>Switzerland</td>
</tr>
<tr>
<td>Brian Ganson</td>
<td>Senior Fellow</td>
<td>Center for Emerging Market Enterprises, The Fletcher School, Tufts University</td>
<td>USA</td>
</tr>
<tr>
<td>Lawrence Lachmansingh</td>
<td>Development Consultant</td>
<td>Independent</td>
<td>Guyana</td>
</tr>
<tr>
<td>Baht Latumbo</td>
<td>Deputy Director</td>
<td>Centre for Peace and Conflict Studies</td>
<td>Cambodia</td>
</tr>
<tr>
<td>Martin Ogango</td>
<td>Consultant</td>
<td>Independent</td>
<td>Kenya</td>
</tr>
<tr>
<td>Meas Sokeo</td>
<td>Co-Director for Programs</td>
<td>The Alliance for Conflict Transformation</td>
<td>Cambodia</td>
</tr>
<tr>
<td>Gopar Tapkida</td>
<td>Regional Peace Advisor</td>
<td>Mennonite Central Committee, West &amp; Central Africa</td>
<td>Nigeria</td>
</tr>
<tr>
<td>Koenraad Van Brabant</td>
<td>Head of Reflective Practice and Learning</td>
<td>Interpeace</td>
<td>Switzerland</td>
</tr>
<tr>
<td>Sue Williams</td>
<td>Director</td>
<td>Summer Peacebuilding Institute, Eastern Mennonite University</td>
<td>USA</td>
</tr>
</tbody>
</table>
Appendix C: Evaluations and Related Documents Reviewed


CDA Collaborative Learning Projects. *Assessment of BCPR-Supported Conflict Prevention Initiatives in Ecuador, Preliminary Notes from Diana Chigas.* CDA Collaborative Learning Projects, N/D.


GTZ. *Guidelines on evaluating the success of projects/programmes in the evaluation priority sector “Peace Building and Crisis Prevention”,* N/D.


No author. *An evaluation of Conciliation Resources, Terms of Reference*. Location unknown, N/D.


Reimann, Dr. Cordula. *Nonviolent Peaceforce Evaluation of NP’s Project in Mindanao, Philippines*. Swisspeace, N/D.


Tapkida, Gopar. *RPP for Evaluation*. Mennonite Central Committee, N/D.

Woodrow, Peter and Chigas, D. *Connecting the Dots: evaluating whether and how programmes address conflict systems*. Systems Thinking and the Evaluation of Peacebuilding Programs, N/D.


Appendix D: Author Biography

Cheyanne Scharbatke Church is the founder of Besa Consulting: Catalysing Strategic Change, a boutique consulting group dedicated to implementing quality change processes. As a practitioner-scholar, Cheyanne’s work revolves primarily around issues of accountability and power. She has conducted evaluations ‘in’ and ‘on’ conflict issues as well as advising agencies on policies and systems to support quality design, monitoring and evaluation systems. She has been a frequent presenter at the American Evaluation Association annual conference and speaks on issues of evaluation in peacebuilding.

Cheyanne teaches classes on evaluation, learning and corruption at the Fletcher School, Tufts University and has specific geographic expertise in West Africa, the Balkans and Northern Ireland. She has held positions with the Reflecting on Peace Practice (RPP), Search for Common Ground and INCORE and consulted for organisations such IDRC, ICTJ, ICRC, UNPBF, and CARE. Her recent publications include Designing for Results, a practitioner focused manual on design, monitoring and evaluation for peacebuilding co-authored with Mark M Rogers; NGOS at the Table: Strategies for Influencing Policy in Areas of Conflict which she co-edited with Professor Mari Fitzduff and Mind the Gap – Policy Development and Research on Conflict Issues.

Besa Consulting: Catalyzing Strategic Change
1-403-460-1951
Cheyanne@besacsc.org