

**The Network for Peacebuilding’s First Thursday Talk on DME for Peace:
Colleen Duggan, Gaps in Peacebuilding Evaluation**



Thank you to all those who joined us for the Network for Peacebuilding Evaluation’s first Thursday Talk on DME for Peace!

Listen to the recording of the session in the video below, or read the written remarks prepared by our speaker, Colleen Duggan of the International Development Research Centre.



A note to those who listened to the call: the remarks below include Ms. Duggan’s thoughts on a third gap in peacebuilding evaluation that she did not have time to include in the live session. To see those remarks, scroll down to [anchor] The Third Challenge: Inexistent or Insufficient Baseline Assessments.



Remarks by Colleen Duggan, Senior Program Specialist at the International Development Research Centre (IDRC)

Introduction



The points I’m going to make today come from learning and experience emerging from the world of evaluation research and practice. Some of this learning is particular to peacebuilding as a distinct area of research and practice. I think the US Institute of Peace (USIP) – Alliance for Peacebuilding (AfP) Peacebuilding Evaluation Project that brought together funders, practitioners and evaluators was quite successful in highlighting the particular challenges of peacebuilding evaluation; similar learning and reflections can also be found in the OECD DAC’s Guidance on Evaluating in Settings of Violent Conflict and Fragility.

But much of the learning also comes from other sectors of that wrestle with providing solutions to complex social problems on a global scale – global health, climate change and adaptation, prevention of HIV transmission, and elimination of human trafficking, to name a few. When it comes to peacebuilding evaluation, I think it’s important to look outside of our own backyard – into the yard of neighbours who are struggling with similar methodological and practical problems. Time is short, so I’m going to organize my remarks around 2 key areas of concern that I consider to be ongoing evaluation challenges or gaps for peacebuilding scholars and practitioners, and if there’s time, I’ll address a third key area of concern as well.



FIRST CHALLENGE: Are we learning the right things at the right time?

I think I’m not alone in my concern that peacebuilding programming has historically been an area of international aid that is particularly vulnerable to the interests of donors, many of whom are located in North American and European capitals. This donor landscape, however, has been changing quite dramatically over the last 10 years and there are a plethora of new funding actors working for “social betterment impacts” – private and corporate philanthropy, social entrepreneurs, micro-foundations, mega-foundations from the Global North and South, particularly in Emerging Economies – in countries like India, China and Brazil. This is fundamentally changing the face and nature of international aid.

Despite this changing landscape, I wonder (and I throw this out as a question to you all): Is the peacebuilding practitioner and scholar community still largely stuck in traditional models of external evaluation? You all know the model I'm talking about: At start-up, a project typically puts together a fairly linear results framework, sets up a monitoring regime (which may or may not be robust); and contemplates a mid-term formative and/or final summative evaluation by a group of externally hired evaluators.



Peacebuilding programming, as we all know, tends to take place in complex and often unstable, unpredictable, fluid settings. It involves working in contexts of social tension with multiple kinds of actors – sometimes former adversaries. And it's almost always deeply political. These settings in many ways represent the antithesis of the methodologically desirable evaluation environment. The very nature of peacebuilding intervention contexts demand that we learn new and different things and that we learn faster in order to correct our course, change or adapt our program design. For peacebuilding, the case for traditional end-of-project summative evaluation is becoming harder to make, and some would argue that it should go the way of the dinosaur.



The first important gap I'd like to mention: We need to get better at using data and evidence in new ways to find answers that inform the decisions we need to make, faster. We need to break out of the traditional evaluation approach that tends to still dominate peacebuilding evaluation. There are three encouraging trends and strategies for dealing with this gap that I'd like to highlight:



MercyCorps

1) "Dynamic monitoring regimes"

There's increasing discussion in international development evaluation circles about "real-time monitoring" – I prefer to call this dynamic monitoring.

What's clear is that monitoring is now about much more than simply collecting data to prove that project activities have been implemented. In response to the dynamic environments in which peacebuilding interventions tend to take place, there's a growing demand for more nimble and flexible monitoring regimes with faster feedback loops. For me, "dynamic monitoring" involves a number of things: (1) a more selective and discriminating approach to how much and what sort of data is collected and with what frequency; (2) an expanded definition of and outreach to stakeholders, increasing their participation in project implementation through new means; and (3) a streamlined approach to data analysis and an accompanying strong commitment to making sense of data.



The evolution, expansion and accessibility of information communication technology is also revolutionizing monitoring practices. Monitoring now has wings. Open access data systems and practices such as data-mining, crowd-sourcing and mobile data collection are casting a new light on how data can be used for faster learning so that organisations can learn, adapt and get better at what they do.

2) Developmental evaluation

Michael Quinn Patton is the primary expert on Developmental Evaluation (DE), though there are others, and DE is gaining a lot of traction in development and peacebuilding communities. DE is emerging as an alternative to traditional model testing evaluation. It positions the evaluator as a part of a programme's design and development process. The evaluator collects information and provides feedback to continually improve the programme. DE is a sort of 'embedded evaluation'

that is done in real-time; it has a series of short, rapid feedback loops, and allows the organization to quickly adapt its strategies and activities.

But developmental evaluation takes time and a certain level of financial investment. Funders and peacebuilding implementers alike need to be convinced that it will pay off in the long run.



3) Increasing movement towards more internal evaluation

The aid industry has seen an increasing move towards internal evaluation in recent years. This is in part a side-effect of the growth of the social benefit sectors and the appearance of new actors, as I mentioned. Interestingly, internal evaluators in these new actor organisations don't always have evaluation training and don't always self-identify as evaluators. While many observers worry about the "objectivity" of internal evaluation, it offers particular benefits to organisations that are serious about organisational learning and closing the gaps between strategic planning, program design, performance management and evaluation.



This links to the point I made about dynamic monitoring regimes. The peacebuilding community needs to invest more resources in strengthening organisational learning - capacities, not only for collecting data; but also capacities for strategic analysis of that data and ongoing reflective practice.

SECOND CHALLENGE: Ethical Evaluation Practice Gap



Evaluation in conflict affected contexts often involves significant risks and high stakes for evaluation actors. Serious consequences – political, security and otherwise – can and do unfold from unethical evaluation practice.



My work over the years as a researcher, a funder and as a commissioner of evaluation has highlighted what I consider to be one of the most under-researched and under-served gaps in evaluation theory and practice: How contexts of violence and conflict influence and inhibit ethical evaluation practice.

In my evaluation capacity-building work, I find that commissioners, evaluators and evaluation stakeholders consistently ask for practical advice on how to deal with ethics dilemmas. Ethics continues to be the missing ingredient in how donors and academic institutions see and engage in building evaluation as a professional field of theory and practice in the Global South.

A brief review of most of the evaluation capacity-building initiatives being offered in the South by major funders of international development & peacebuilding indicates that these focus almost exclusively on the technical/methodological side of evaluation – with little discussion of the relational dimensions of evaluation and without any serious focus on the rights and responsibilities of both evaluators and commissioners/managers. This could be a result of outmoded approaches to training and technological transfer - but I suspect that has more to do with two things:

(1) The fact that ethics in evaluation is slippery and hard to pin down. What constitutes an ethical conflict is not always clear; one evaluator's ethical issue might be an evaluation commissioner's political or methodological problem (see Michael Morris's book *Evaluation Ethics for Best Practice: Cases and Commentaries*, 2008).

(2) Both evaluators and commissioners of evaluation often have moral, practical and political reasons for wanting to frame up ethics conflicts as technical or methodological problems. This

becomes “extreme” in conflict affected settings – places in which there is a high premium placed on donor visibility and there is acute pressure to be seen to be “doing good” and not exacerbating conflict (particularly in peacebuilding interventions).



I am also struck by the numerous examples of ethics guidelines and principles for evaluators – with very few equivalent “enforceable” or binding standards for the commissioners and funders or evaluation. I am aware that this issue of power imbalance between the evaluator and his/her client has occupied the attention of evaluators for years. Yet, if evaluation is a form of research, at what point should an evaluation design and the ongoing conduct of an evaluation undergo some sort of external ethics review and oversight? Universities have ethics review boards to address exactly this issue.



Arguably, evaluation advisory groups have been known to play this role. Even so, there is a curious grey zone between ethics guidelines for evaluators, evaluation advisory groups, and contractual obligations/requirements for evaluators. In my own organisation, the contractual provisions we have with evaluators mean that they have no intellectual property rights over their evaluation data or reports themselves.



Recognising, planning for and grappling with ethics challenges in conflict-affected settings is not easy and does not easily lend itself to standardization or operationalization, much less institutionalization.

THIRD CHALLENGE: Inexistent or Insufficient Baseline Assessments



MercyCorps

Linking baseline assessments with contextual analysis and conflict analysis is a gap that many are aware of, but are still struggling to address.

The absence of baseline data in the contexts in which peacebuilding interventions typically take place is a problem that is consistently cited by peacebuilding practitioners, funders, scholars and evaluators. I think this problem can actually be parsed out into several layers:



1) Collecting baseline data in conflict-affected areas

It’s true that baseline data is often difficult to collect in many peacebuilding programming sites. There are many reasons for this: knowledge infrastructure has been damaged or destroyed; entire populations or demographic make-up have shifted; or project designers may have lacked the technical capacity to undertake a baseline assessment.

2) Baseline assessments practice gap

My sense is that although baseline assessments are part of good research practice, sadly, they do not happen as consistently as they should in the design of peacebuilding projects and programs. And this is a problem because no matter how painstakingly a Theory of Change has been developed at program outset (and adapted as a program evolves), it is of limited use if we have no sense of what the situation looked like before project start-up. Any tracking of progress forward or measurement of results achievement sits on a shaky methodological foundation in the absence of reliable data.

Recognizing this dilemma, researchers and evaluators have come up with strategies for re-constructing baselines: using secondary data from governments, NGOs or donors; using project administrative data (needs assessments); and/or using recall techniques. These often need to be adapted for the special circumstances that often characterized settings affected by conflict and fragility.

3) Incorporating conflict analysis

There's been more discussion and a growing body of literature on the incorporation of conflict analysis into project and program design. This is good. But evaluators are wrestling with how they should fit conflict analysis into the evaluation design, and are also grappling with how conflict analysis (depending on whose analysis this is - because this also matters) interfaces with baseline assessment. This is a tricky methodological gap – and I'd be interested in hearing about experiences and experiments that others in the peacebuilding community are undertaking to address this gap. Where is the practice moving?



Question – Answer Period, Including Questions for Further Discussion



We had so many great questions from participants in the Thursday Talk! The questions participants submitted are listed below, and organized into five categories: 1) Evaluation ethics; 2) Evaluation methodologies; 3) Resource requirements for evaluation; 4) Donor policies and practices in evaluation; and 5) Examples of best practices. Ms. Duggan's remarks are included below the appropriate question, but unfortunately due to time constraints, she was not able to respond to all questions. We encourage all of you in the community to respond with your own insights, thoughts, lessons learned, and experiences. Let's keep this conversation going!



1) Evaluation Ethics

Ethical review and the evaluator – commissioner relationship



MercyCorps

Q: How is the issue of "accountability imbalance" in peacebuilding relevant to you as a peacebuilding evaluator?

Q: How do ethical practices in evaluation differ from research?



Q: I'm struck by the point about evaluation advisory boards, as an interesting mechanism to regulate the grey zone that you described that evaluators face in terms of their ethical challenges and concerns. Could you expand a bit, what does that mean- is that something for the field, or for the organization to handle? In universities, we have ethics boards and other mechanisms to address these challenges.

Colleen Duggan: One thing that sets the ethics evaluation boards that you find in universities apart is that they are fairly well structured, there are points at which researchers must submit their evaluation design, and depending on what kind of methodological, ethical concerns the board might have, researchers are expected to check in at certain times.

Theoretically, this could happen, in the context of an evaluation and evaluation design, but often, it doesn't happen, and it's actually quite surprising, because, as I tried to highlight, working in conflict-affected settings and on peacebuilding projects and programs – which often in many ways have quite strong political objectives behind them, different sets of expectations by the actors that are involved- that there are not more "check in points" or more standard channels of communications, in which evaluators and their clients/commissioners or managers of evaluation, as I call them, can actually discuss some of the ethical questions that come up. For example, a very typical example that I can think of, is evaluators who are working in conflict-affected settings, in the context of their data collection, can stumble across things they might not know about, such as instances of wrongdoing, such as human rights abuses that have happened, and this can come across in the process of interviewing stakeholders, or any place in the data collection. So the

evaluator has to decide what they do with that information; they may feel a strong moral obligation to be forthcoming with that information, but on the other hand, they are somewhat bound by the confidentiality understandings that researchers have with their research subjects, and that's especially acute if, going into an evaluation, an evaluator has said at the outset that the information the interviewee gives will be held in confidence. And it's quite hard to know what to do in those situations.



There are other kinds of dilemmas which come up, which are more the product of external factors bumping up against what might be an established ethical principle. For example, one assumes that in the context of an external evaluation, funders or commissioners of evaluation hire an external evaluation team because they want some independence or objectivity in the evaluation exercise. Yet, we often find that funders or commissioners get upset by what evaluators come back with, if the findings aren't in line with what the funder/commissioner expected, and there's often pressure applied on evaluators to change their findings. And that can often become an ethical dilemma; and though some people may say that's just political, it can be ethical. For example, if after an evaluator has presented a report, the commissioner extinguishes the voice of key stakeholder groups, or calls into question the opinions and perspectives that have been put forth by key stakeholder groups, then that becomes an ethical dilemma. There have to be spaces in which evaluators can discuss this, both with the people they're evaluating, but also with the people who have commissioned them to do the evaluation.



MercyCorps

2) Evaluation Methodologies

Internal vs. external evaluations

Q: What is the difference between internal and external evaluations, and what measures need to be in place to make the internal evaluator useful?



Q: What are your comments on the stress between internal and independent evaluations?
Colleen Duggan: I'm not by any means advocating strictly internal evaluations for all evaluation work. As background, internal evaluation often refers to an evaluation that is done within an organization, and there are different ways it can be done, from self-assessment exercises to evaluation units who manage evaluations. So the objectivity issue is certainly a big issue, especially in the domain of peacebuilding, as there are often concerns that because people on the ground often have many vested interests, both NGOs themselves, government actors, different community groups/factions, it's important to have an external, objective perspective, which is one of the reasons external evaluators are hired. However, my sense is that often in the peacebuilding area, evaluators are hired who are from distant capitals, and they don't always have the requisite competencies; sufficient grounding in understanding local dynamics or even the history of the conflict; and this is a problem with external evaluation. However, that said, there are often reasons why an external, objective perspective needs to be brought in. This is often the case when a project or programme has quite lofty claims of achievement; there will always be doubts if a program claims quite lofty results or goals, so it is always useful to have external evaluation to support that. So in the organization I work for, we do a combination of internal and external evaluations. But one of the big benefits of internal evaluation is organizational learning. NGOs, practitioners and implementers have to get much better at understanding what effectiveness is for them, to make their programs better. So I'm a big fan of internal evaluation, and I've noticed over the years that

there is a growing movement towards internal evaluation, but it's definitely an area that needs more work.

Dynamic monitoring regimes



Q: Can you please elaborate on how to expand effective participation of stakeholders in a "dynamic monitoring regime"?

Colleen Duggan: I think that expanding the participation of stakeholders in monitoring activities really comes down to a case-by-case basis. It depends on what a project looks like, but I think that peacebuilding practitioners should try to actually carve out time to have sustained conversations with their stakeholders in the lifetime of the project. It's quite easy to get bogged down and talk to stakeholders about how project activities are going; are we on schedule, when's the next meeting, etc. But I think there are opportunities for reflective learning sessions, more types of activities like focus group discussions that we can do with the people we're working with on the ground. There's a lot of literature out there, particularly SAS2 Social Analysis Systems: A Guide to Collaborative Inquiry and Social Engagement by Jacques Chevalier and Daniel Buckles, which actually look at things like participatory statistics and methods for undertaking what I would call dynamic monitoring.



Integrated indicators



MercyCorps

Q: Do you need integrated indicators for evaluating impact of peacebuilding work in fragile states?

Colleen Duggan: I used to be strongly against standardized indicators, but the difficulty I find with standardized indicators is that people often get lazy, and they'll just look to a list of indicators and replicate those across contexts and in different circumstances without thinking whether this is a meaningful indicator for my project or programme. That said, I think everybody, when they're designing a project or programme, are always struggling for ideas about what might meaningful indicators look like. So I think that the movement towards having large banks and repositories of indicators and the extent to which people are exchanging information is useful, but I do have some concerns and doubts regarding the application of standardized indicators across different contexts, spaces, circumstances, etc.



Integrating evaluation through a project, programme, organization, etc.

Q: How is project evaluation nested within larger organizational learning incentives and structures, or how should it be?

Colleen Duggan: There is a sentiment that is often drudged up with the word "evaluation," which for many people working in international aid, including peacebuilding, strikes fear into their hearts. It strikes notes of people feeling like external evaluation is often inefficient; it's a waste of time; evaluators tell us what we already knew, etc. So there is a stronger movement towards what I would call embedding evaluative thinking into programming and program design, so that people use data and evidence to consistently and constantly question their assumptions, and the assumptions that underpin programs. I do think that Theories of Change (TOCs) can be useful; it's helpful to set out a TOC during program design, but TOCs are only good if they are continually interrogated and probed, and you need data and evidence to do that. So I thoroughly agree that it is critical for peacebuilding implementers and practitioners to try and strike that balance between getting on with the very important work they're doing, but also finding a way to use the data that comes in for actual strategic learning and adaptation. I think, as we move forward, people are getting more and more comfortable with not placing labels on these terms. What interests me, is

that people are using data and evidence to get better at what they're doing, regardless of what they call it, but as long as it is built in their program, that's really important.



In project design, there has to be a conversation about how results will be articulated. So often, in peacebuilding, evaluations end in "soft results," data that is not easily quantifiable, such as changing relationships between groups of people. As such, there should be more thinking at the start of projects about what success would look like, what progress would look like, and this thinking should be used to guide the evaluation process.



Participant remark: The other issue with results, as well as unquantifiability, is the impact of the programme alongside other factors and programmes

Participant remark: I agree, and I wanted to summarize, regarding the need for integrating evaluation into projects/programs. I think this gets to the relational dimension spoken of. Inclusive relationship-building is critical in fostering the commitment needed for "effectiveness" and collaboration among donors, implementers, and evaluators, and even independent external evaluators. After all, the central distinguishing dimension of collaboration and sustainable partnership (for peacebuilding and any other cross-sector development initiative) is inclusive participation for identification and mediation (and certainly resolution) of conflict. (In this case, conflict refers to the ethical conflict that has been spoken of in this meeting).



MercyCorps

3) Resource Requirements for Evaluation

Maintaining commitment to evaluation

Q: As mentioned in the talk, implementers and donors often shy away from developmental evaluation because it requires more time and resources (which are always short during implementation). At the planning stages and after projects, the drive for better evaluation is always high, but during the program implementation and wrap up evaluation, this drive can fall by the wayside. Do you have any recommendations for ways to address this challenge and maintain commitment to evaluation throughout the project life cycle?



Q: Can we frame dynamic evaluation as an integral part of the program, rather than an extra expense, and if so, how?

Colleen Duggan: I think it's a huge challenge, maintaining commitment to evaluation. It's interesting where I sit in IDRC, because we're a funder, but we also have other organizations that entrust their resources to us. In the context of those types of partnerships, it often comes down to resources; how much of the budget can we be using for M&E? There's often pressure to use as much of the resources as possible on the actual project. That's why I come back to the issue of monitoring, or even farther back, when you get to project design, I think you can gain a lot of ground by actually investing time up-front and on having a wide discussion on how results are going to be articulated. That is particularly important in peacebuilding, because often results are intangible, hard to pin down; they're often "soft" results, involving reconstruction of relationships, things like reconciliation or transitional justice. They're things that aren't easily quantifiable. So I think that there needs to be more evaluative thinking, more strategic thinking put in at the front end of the program cycle about what is it that we're actually trying to achieve, and what will success look like? What will we consider to be an acceptable result? What will progress along the road look like? That's where monitoring can really come in. I think developmental evaluation (DE) is, in many ways, an interesting innovation, but in many ways, I think that there are elements of DE

that, to me, speak strongly of internal evaluation. The difference with DE is that the evaluator acts as a critical friend. So, I think that implementers, project implementers and those of us involved in aid, are very caught up in the day-to-day hustle of getting on with the project. It's useful to have someone who is a bit of a critical friend, standing back and helping make sense of monitoring data, and helping project implementers ensure that there is space for reflection. This all sounds excellent in theory, and I am keenly aware of the constraints, but I really think it's what we need to be pushing towards. This was a big push behind things like outcome mapping, was actually about ring-fencing some important space for reflective practice and continual thinking about how we're adapting and marking results along the way.



Allocating resources for evaluation

Q: From a donor's perspective, can you suggest what percentage of a peacebuilding programme should be earmarked for evaluation?

Q: If donors require evaluation of peacebuilding programmes they fund, how do they decide how much of their grant should be spent on evaluation? Is there a minimum or maximum? What criteria should be used?



Colleen Duggan: I don't think there is an easy answer to what percentage of a budget should be earmarked. And maybe this goes back to my roots in utilization-focused evaluation- I don't think evaluation should be undertaken unless there's a clear reason why an evaluation, especially if it's



MercyCorps

an external evaluation, is being commissioned, unless there is clear support that the evaluation is going to be used, and unless it's clear who the user is going to be, because evaluations can be financially quite costly, and take people's time. I don't think it's helpful to earmark x percentage for evaluation of peacebuilding. I think the question is, what is your project about? If your project or programme is breaking new ground, undertaking innovative new work that no one has tried



before- then that, to me, says evaluation maybe should be part and parcel of what you're doing. So it comes down to undertaking evaluation when you think there's a solid rationale for doing it, and whether you think there's a solid commitment for actually using the evaluation. We need to move away from a very formulaic approach of just tacking evaluation onto programs.

4) Donor Policies & Practices in Evaluation

Ethical guidelines for donors

Q: Do you see much movement in the donor community to set ethical guidelines?

Donor practices

Q: I think we are hemmed in by the word evaluation. What we are talking about is learning, and this has impacts throughout an organization- funding, finance, personnel, etc. How can donors reward good practice here?

Participant remark: We seem to be talking about Design, Monitoring & Evaluation (DME) as a continuous process, internally managed/led with expert support from external evaluator(s). There seems to be progress in getting donors to recognise and support that.

5) Best Practices Examples

Q: What are some concrete examples of good practices, with or without the names of the organizations?

Q: What is an example of what you consider a good, low-cost not-too-time-intensive evaluation in peacebuilding?