PLEASE SCROLL DOWN FOR ARTICLE

EDITORIAL

EVALUATION IN VIOLENTLY DIVIDED SOCIETIES: POLITICS, ETHICS AND METHODS

Those who work in support of peacebuilding and development initiatives are acutely aware that conflict-affected environments are volatile, unpredictable and fast-changing. In light of this reality, evaluation and research in the service of peacebuilding and development is a complex enterprise. Theories of change and assumptions about how peace and development work are often unarticulated or untested. While much work continues to be done on the theories, methodologies and praxis of peacebuilding, we suggest that the international aid community, researchers and practitioners need to think more deeply and systematically about the role of evaluation in increasing the efficacy of projects and programmes in violently divided societies (VDS).

Core questions that underpin and motivate the articles contained in this special issue of the journal include:

- How does the particular context of conflict affect our approaches to, and conduct of, research and evaluation?

- Specifically, how do politics — be they local, national, international, geopolitical — interact with evaluation practice in ways that enhance or inhibit prospects for peace and sustainable development?

- What can we learn from current research and evaluation practice in the global North and South about their impacts in violently divided societies?

- Which tools are most effective and appropriate for assessing the role of context? Should there be generic or global assessment frameworks, criteria and indicators to guide evaluation in violently divided societies, and, if so, what do they look like? Or does the fluidity and heterogeneity of different conflict zones inhibit such developments?

- How can evaluation, in its own right, catalyse positive political and societal change? What theories of peacebuilding and social change should best guide evaluation research and practice in ways that promote peace and sustainable development?

The inspiration for this special issue is rooted in a four-year collaborative research project between the Evaluation Unit of the International Development Research Centre of Canada and the International Conflict Research Institute (INCORE) of the University of Ulster. That project highlighted and explored an important but neglected area of research and practice: the evaluation of the impact of *research* on, and in, violently divided societies.[1] The current issue expands its focus beyond *research* to include the evaluation of different types of interventions in conflict contexts (broadly defined).

The *Journal of Peacebuilding & Development*, based at the Joan B. Kroc School of Peace Studies, is a natural partner and host for an exploration of the questions posed above.

1

Since its inception, the journal has been casting a critical eye on the intersections of the politics, methods and ethics of peacebuilding and development. The journal seeks to honour its mission to nurture and disseminate empirically based and contextually rich studies, and to actively seek out, encourage and publish authors from the Global South — in particular, from violently divided societies. The journal publishes a great deal on issues of design, monitoring and evaluation (DM&E) of peacebuilding and development. This emphasis is premised on the belief that good DM&E is the sine qua non of good practice, which, in turn should inform decision-making, and generate better policy. Less appreciated is the fact that robust DM&E shares an iterative relationship with good theorising, both normative and evidence-based — particularly from a bottom-up and/or grounded theory perspective. The achievement of better results in peacebuilding and development demands an ongoing and critical assessment of what works (or not) and why it works (or not). Only then will we be able to generate the empirically grounded, theoretically informed, knowledge required as the foundation for good scholarship, policy practice and activism.

When examining the implications of conflict context for theory and practice, we are considering both (1) the influence of violence or conflict on the environment — physical, historical, social, cultural, political, organisational — within which evaluation practice occurs, and (2) the ways in which the very existence of violence (its presence, legacy or potentiality) may influence how stakeholders engage (or not) with projects, programmes, research or evaluations. This includes the consideration of how context affects prospects for communication, uptake and actual use of evaluations for social change. Our contention is that in settings affected by significant levels of militarised or non-militarised violence, context is much more than a landscape or backdrop. It is a factor that permeates and affects all aspects of an intervention. Further, within such conditions, even apparently insignificant actions may serve to ignite a powder keg of latent conflict or violence. It is the very existence of violence and the particular conditions that it spawns and superimposes upon the context which raise acute challenges for development workers, peacebuilders, researchers and evaluators.

Important progress has been made in the areas of peace and conflict impact assessment (PCIA), conflict-sensitive development methods, and peacebuilding evaluation. After years of research, testing and reflection the OECD DAC[2] in November 2012 released its guidance on 'Evaluating Peacebuilding Activities in Settings of Conflict and Fragility'. It urges donors and partners to base their work on a clear understanding of conflict context and to develop clear, evidence-based theories of change. New initiatives of consortia[3] have been formed to document, systematise and advance the field of 'peacebuilding evaluation' towards a more professional and systematic area of research and practice. These include a portal of peacebuilding evaluation that offers over 60 different peacebuilding evaluations for researchers and practitioners to examine, alongside a meta-evaluation study outlining lessons across the field. As well, the Peacebuilding and Evaluation Consortium (PEC) was launched in 2013; it aims to build capacity in the field through various means: developing a network for practitioners and academicians to provide further support to evaluators, the development of guidance and manuals on how to carry out systematic evaluation, and sharing of case studies to publicise lessons learned from various cases of actual evaluation.

At the policy level, the International Dialogue on Peacebuilding and Statebuilding, with its drivers — the g7+, the International Network on Conflict and Fragility

(INCAF) and the Civil Society Platform on Peacebuilding and Statebuilding — is engaging the demand for appropriate goals, frameworks and methods to assess countries' fragility and progress towards peace. Concurrently, the heated debates around what belongs in the post-2015 Development Agenda are being strongly engaged by the peacebuilding community to ensure that all development practice engages with issues of conflict, violence and fragility. The *Journal of Peacebuilding & Development* monitors and reports on these intense global debates in its 'Policy Dialogues' section.

The current issue of the journal is an important, if modest, contribution to these peacebuilding, development and evaluation initiatives. While many of the above noted initiatives focus on methods to evaluating peacebuilding initiatives, our analytical ambit includes all types of interventions — developmental, humanitarian, peace-building, private sector investment and so on. This allows us to emphasise and explore the impact of different conflict contexts on interventions. It also allows us to explicitly include all the thousands of *other* projects and programmes in conflict zones (i.e. in more traditional development related areas) that are excluded from the narrow peacebuilding lens but nonetheless have profound peace and conflict impacts. Additionally, we do not restrict our case selection to militarised zones of violent conflict. Analytically, we include cases from the broadest spectrum of conflict environments — be they characterised by 'proto-violence', social violence, criminalised violence or militarised violence. This provides a richer source of potential learning around the conflict and peace related interactions of programmatic work within such contexts, providing opportunities to better identify and understand when, why and how non-violent conflict *becomes* violent. By bringing this broad peace and conflict impact lens to the evaluation of *all kinds of research and programming*, our intention is to bridge discussions that are taking place in different literatures, fora and areas practice — which have a tendency to be siloed.

It is our hope that this special issue will contribute to current debates and efforts in a number of ways:

- Deepen dialogue around these issues that challenge the fields of development and peacebuilding and inhibit our work.

- Focus critical attention on issues that need to inform efforts to improve the DM&E of interventions in conflict contexts — in particular, politics, logistics, methods and ethics.

- Highlight the need to better understand the impact of conflict context on evaluation — and the impact of the evaluation on the (peace and) conflict context.

- Increase awareness that conflict context forces us to think about, and conduct, our evaluations differently; and that efforts to employ or to impose evaluation-as-usual evaluations (i.e., non-conflict sensitive) in conflict zones risks inflicting profound harm.

- Underscore the need for a broader dialogue between evaluation research and peace and conflict studies.

Kenneth Bush and Colleen Duggan (Visiting Editors), and Erin McCandless and Mohammed Abu-Nimer (Journal Issue Editors)

Endnotes

[1] Findings from this project will be published in a collected volume, "*Evaluation in Extremis. Research, Impact and Politics in Violently Divided Societies*" (forthcoming, 2013)

[2] Organisation for Economic Co-operation and Development, Development Assistance Committee

[3] The portal involves the Alliance for Peacebuilding (AFP), Search for Common Ground (SCG) and the Center for Peacebuilding and Development (CPD). The PEC includes AFP, CPD, SCG, Mercy Corps and the United States Institute for Peace (USIP), and is supported by the Carnegie Foundation.

Evaluation in Conflict Zones: Methodological and Ethical Challenges

Kenneth Bush & Colleen Duggan
Published online: 19 Sep 2013.

PLEASE SCROLL DOWN FOR ARTICLE

ARTICLES

EVALUATION IN CONFLICT ZONES: METHODOLOGICAL AND ETHICAL CHALLENGES

KENNETH BUSH AND COLLEEN DUGGAN

Abstract

This article explores the methodological and ethical challenges particular to the conduct and use of evaluations in conflict zones. It does this through examining the synergistic interaction of conflict dynamics and the four domains of evaluation — ethics, methodology, logistics, and politics. Drawing on evaluation theory and practice, as well as field experience, the article seeks to contribute to the building of a more methodologically self-conscious sub-field of evaluation in conflict zones — with implications not only for the field of evaluation, but also for researchers and practitioners in the fields of development, humanitarianism, peacebuilding, and private sector investment.

Keywords: evaluation, methodology, ethics, politics, conflict zones, fieldwork, conflict theory, peacebuilding, development, complexity

Introduction

This article explores the methodological and ethical challenges particular to the conduct of evaluations in conflict zones. More specifically, it asks: how can (or should)

> How can we improve evaluation practice to better understand the difference, whether positive or negative, that interventions make in societies divided by violent conflict?

conflict context affect the way we commission, conduct, disseminate, and use evaluations? And how can we improve evaluation practice to better understand the difference, whether positive or negative, that interventions make in societies divided by violent conflict?

The need for such an exploration is two-fold. First, the proliferation of international initiatives in violence-prone settings following the end of the Cold War has created a commensurate increase in the need to evaluate them in ways that are sound and appropriate methodologically, politically, and ethically. This applies to private sector investments no less than not-for-profit humanitarian, development, or peacebuilding

5

initiatives. The second, and somewhat more puzzling, issue to which this article responds is the need to redress the relative absence of the systematic consideration and incorporation of conflict context into the theory and practice of evaluation (Organisation for Economic Cooperation and Development [OECD] 2012, 19).[1] The approach in conflict zones is essentially evaluation-as-usual. The problem is that the working environment is far from 'usual'. Conflict zones, in many ways, represent the antithesis of the methodologically desirable evaluation environment: they are unstable rather than stable; unpredictable rather than predictable; and far less 'controllable' than non-conflict evaluation environments.

> **Conflict zones, in many ways, represent the antithesis of the methodologically desirable evaluation environment: they are unstable rather than stable; unpredictable rather than predictable; and far less 'controllable' than non-conflict evaluation environments.**

This article acknowledges and builds on earlier work undertaken on peace and conflict impact assessment (PCIA),[2] as well as the subsequent methodologies and initiatives that spun out from that original research.[3] It also acknowledges the important policy work by the OECD on the evaluation of 'conflict prevention and peacebuilding activities' (OECD 2007; 2008; 2011; 2012). However, the focus of the article differs from these earlier (and ongoing) initiatives in a number of important ways: (1) contra the OECD emphasis, we focus on the evaluation of *all forms of* interventions in conflict zones, not only those which are self-labelled and funded as 'peacebuilding and conflict prevention' initiatives; (2) we define 'conflict zone' more broadly than areas of militarised violence (as addressed below); and (3) we attempt to more explicitly integrate evaluation research into the existing emphases on policy and peacebuilding programming.

This article draws on research and practice from a number of initiatives rooted in four years of collaboration between the Evaluation Unit of the International Development Research Centre (IDRC, Canada) and International Conflict Research (INCORE) at the University of Ulster in Northern Ireland; specifically: (1) a research project that brought together researchers, evaluators, and funders who commission evaluations — three groups of individuals from the global North and South, who share a stake in the improvement of evaluation and research practice in conflict-affected settings; (2) INCORE's annual summer school course on evaluation in conflict-affected settings, which attracts individuals from international development and humanitarian agencies, NGOs, evaluators, and students and scholars of peace and conflict studies; and (3) ethics training workshops for evaluators which the authors have developed and delivered in Africa and South Asia.

The article is structured as follows. The first part presents an introductory discussion of the nature and implications of conflict context for evaluation. The second part introduces and examines four core domains of evaluation. Each of the next four sections explores the ways each of these domains — logistics, politics, methodology, ethics — interacts with conflict context. Because the primary focus of the paper is methodology and ethics, more emphasis is placed on the latter two domains.

Evaluation in Extremis: What Difference Does Conflict Context Make to Evaluation?

Every conflict zone has its own particularities. But conflicts are neither sui generis nor entirely unique. There are patterns and commonalities within and across conflict zones.

All of them are characterised by varying levels and forms of fluidity, uncertainty, volatility, risk, and insecurity. Such characteristics serve to amplify the impacts of the conflict environment on evaluations, no less than on other forms of intervention. This article seeks to highlight and explore the ways these features affect the conduct and consequences of evaluations in conflict environments.

To be clear, the term 'conflict context' is used in this article to refer to:

- the influence of conflict on environmental conditions (physical, historical, social, cultural, political, organisational) within which an evaluation is undertaken; and

- the impact of conflict (its presence, legacy, or potentiality) on stakeholder interaction with evaluators, their assessments, and prospects for communication and actual use of their findings.

The term 'conflict zone' is *not* limited to areas of militarised violence. It includes sites of non-militarised violence, characterised by: gendered violence; class and caste violence; social violence; political instability; state-sanctioned intimidation; structural violence; (un)organised crime; through to genocidal violence (Scheper-Hughes and Bourgois 2004). This particular understanding of 'conflict zones' carries methodological implications requiring us to consider the broadest range of cases within the scope of our inquiry — from inner city New York to indigenous regions of North America. Such areas are typically characterised by the increased risk, volatility, and potentially damaging consequences that characterise conflict zones.

In settings affected by significant levels of conflict, context is much more than the background for evaluation. Under such conditions, all of the usual challenges to effective development evaluation become more extreme and potentially more debilitating. Thus, for example:

- it is *more difficult* to discern and delineate the specific impacts[4] of an intervention (i.e., attribution problems);

- it is *more difficult* to temporalise a timeline between intervention and impact; and, somewhat counter-intuitively,

- there is a positive bias toward the achievement of outcomes — something that is *more difficult* to detect or challenge because of barriers to follow-up in conflict zones, such as the inability to re-access stakeholders, whether owing to levels of insecurity or the considerably higher economic costs of working in militarised conflict zones.

Choices regarding the planning and conduct of an evaluation in a conflict zone, as well as the use of its findings, are far from technocratic decisions; they are imbued with intensely political, ethical, methodological, and logistical challenges.

The Core Domains of Evaluation

To better understand the interactions between conflict, context, and evaluation, we have conceptually disaggregated evaluation into four constituent components, or 'domains': methods, logistics, politics, and ethics. Figure 1 illustrates the ways in which conflict context impinges on these domains individually and collectively. These domains are nested in a symbiotic relationship with structures and processes of conflict. That is, they

7

will affect, and be affected by, conflict (illustrated here by the two-way arrows between the extreme context of the conflict zone and each domain). For example, conflict context will affect the logistics and methodology of an evaluation; *but* decisions about methodology and logistical arrangements for an evaluation may also affect conflict dynamics. The figure highlights the intersection and interaction of these domains. By framing the major challenges to evaluation in this way, we may move beyond static, one-dimensional checklists of issues within each domain — as in: the methodological challenges are X, the ethical dilemmas are Y, the logistical challenges are Z, and so on. While such lists are necessary, they are insufficient for understanding and exploring the dynamic complexities and synergies affecting evaluation within conflict contexts.

> **Conflict context will affect the logistics and methodology of an evaluation; *but* decisions about methodology and logistical arrangements for an evaluation may also affect conflict dynamics.**

The points of domanial intersection are labelled in Figure 1 (as 'A' to 'E') to correspond to the examples below. These are intended to help us better delineate the kinds of issues that may arise within these intersections.

A. *Ethico-methodological issues* — e.g., reliance on an evaluation methodology that 'disappears' key stakeholders, thereby misrepresenting the impact of a programme and further marginalising an already marginalised group.

B. *Logistico-methodological issues* — e.g., lack of access to stakeholders (owing to time, insecurity, or geography) which compromises methodological integrity.

C. *Politico-logistical issues* — e.g., when an evaluator is *only allowed* to see 'model'/ successful sites by the project implementer.

D. *Ethico-political issues* — e.g., when the client who commissioned the evaluation applies pressure on the evaluator to change the findings, or to write a positive evaluation, when data do not warrant it.

E. *Omni-domanial issues* — e.g., insistence by a client on the exclusive use of randomised control trials, thereby delegitimising all other (context-appropriate) methods that would allow for a more robust evaluation; the motive, in this example, being the desire to cook the results to justify cutting programmes to which the client is ideologically opposed.

Figure 1: The Intersection of Evaluation Domains in Conflict Zones of Moderate Intensity
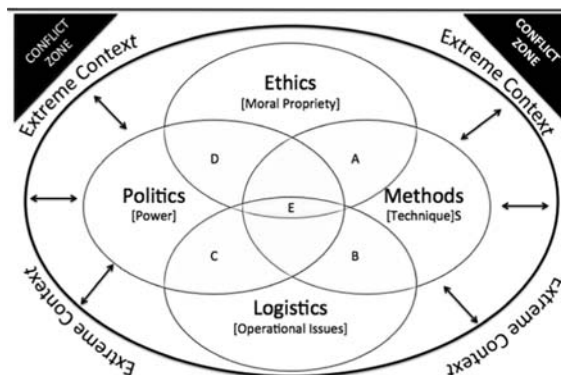
Figure 2: The Amalgamation of Evaluation Domains in Conflict Zones as Intensity Increases
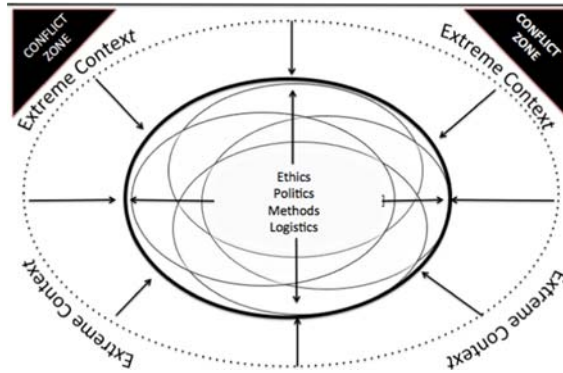
By way of explanation, it may be helpful to compare B (logistico-methodological) with C (politico-logistical). While both examples illustrate a situation where an evaluator is unable to access a sufficiently representative sample of stakeholders as required for a robust evaluation, the reason for the problem is different in each case. In the former case, logistical obstacles (time, geography, or insecurity) affect an evaluator's ability to employ a rigorous enough methodology. In the latter case, vested political interests actively block access to stakeholders — something the evaluator may or may not recognise. These vested interests could be held by any number of actors in a conflict zone — host or home government officials, programme staff of the initiative being evaluated, or the commissioners of an evaluation.

The relationships between the domains and conflict dynamics are fluid as well as interdependent. They may shift over time if conflict intensifies, further constraining an evaluator's latitude of action. As conflict intensifies — i.e., as volatility, risk, and levels of potential harm increase — the four domains of evaluation are forced *into each other* so that decisions and actions in one domain inevitably affect all domains — see Figure 2. Thus, for example, it becomes increasingly difficult, if not impossible, for logistical issues to be addressed independently of ethics, politics, and evaluation method. While this dynamic may also be evident in non-conflict contexts, the difference here is the acute levels of risk and the speed with which relatively minor problems (or miscalculations) in one domain may trigger a chain reaction of serious proportions. This begins to shed light on why evaluation is so much more difficult in contexts affected by violent conflict.

> **The relationships between the domains and conflict dynamics are fluid as well as interdependent.**

Having mapped, in broad brush strokes, how these domains interact and intersect with each other and with conflict context, we now turn our attention to an examination of each of them in turn.

Domain I: Conflict Context and Evaluation Logistics

The logistical challenges to evaluation in conflict zones are manifold. Conflict may affect the safety and security of evaluation stakeholders, affect access to stakeholders or territory, require sudden changes to the design and conduct of an evaluation,

increase levels of suspicion and distrust (thereby hampering data collection), and so on. Such challenges mirror those confronting any fieldworker in the same environment — researchers, development or humanitarian workers, community workers, contractors, and so on. The over-arching, and inescapable, conflict-prone nature of the environment creates conditions of increased risk and decreased predictability — often forcing evaluators to make quick decisions affecting the safety of stakeholders and the feasibility of the evaluation. Although such decisions must typically be undertaken using insufficient or poor-quality information, their consequences are often immediate.

The ways in which conflict context may affect evaluations are similar to its impact on other forms of field research in conflict zones. Since the early 1990s, there has been an increase in the number and variety of publications on the challenges of working in conflict zones, tailored to the specific needs of researchers, development workers, UN workers, peacekeepers, journalists, and so on.[5] These publications sketch out both the nature of the logistical challenges in conflict zones as well as practical responses. The range of topics addressed in this literature is vast, illustrating the diversity and multiplicity of logistical challenges affecting those working in conflict zones: health challenges; insecurity; risk of being kidnapped or raped; emergency preparedness; landmine awareness; media management; and psychological challenges and trauma; to name but a few. The plethora of manuals, guides, and handbooks can describe only the tip of the logistical iceberg.

Domain 2: The Politics of Evaluation in Conflict Zones

Politics perforate evaluation. By 'politics', we are referring to the competition for power in all of its forms from micro to macro levels. At an international level, this includes the foreign policies of states directly involved in the conflict zones where evaluation is undertaken. For example, if the commissioner of an evaluation is an agency of a government against whom insurgents are fighting, then evaluators can expect collateral negativity and risk. Further, to the extent that development assistance is seen by local insurgents to be part of a counter-insurgency strategy, as in Afghanistan, then anyone associated the programme — including evaluators — becomes a target. Even the *domestic* policies of foreign states associated with an evaluation can have negative political implications for evaluation teams. For example, the application of anti-terrorist and anti-immigrant legislation throughout Europe inevitably affects the receptivity to evaluators in the corresponding countries (among both local populations and government officials).

> **Further, to the extent that development assistance is seen by local insurgents to be part of a counter-insurgency strategy, as in Afghanistan, then anyone associated the programme — including evaluators — becomes a target.**

Politics are present in even more subtle forms at the very start of the evaluative process in the interests that animate an evaluation: Who wants the evaluation? How will it be used? What questions will it answer (or not answer)? As Jayawickrama and Strecker (2013) point out, there are even political implications in epistemological and methodological choices in the design of an evaluation; choices that legitimate some voices and realities while delegitimising and disappearing others.

Politics may be evident in different forms at different stages of the evaluation process. At the commissioning stage, they may be embedded in decisions to underfund an

evaluation, so that programme deficiencies do not come to light. At the design phase, politics may be evident in the imposition of an evaluation methodology that inhibits a critical examination of the broader societal impacts of an initiative — or that predetermines the findings of the evaluation — as illustrated in the case where an evaluator turned down a contract because the required methodology would have ensured a positive assessment of the drug trials undertaken by a UN agency but funded by the pharmaceutical company manufacturing the drug.[6] At the data collection stage of an evaluation, politics may be evident in the manipulation of information or access to evaluation stakeholders by those with politically vested interests in the results of the evaluation. So, too, may politics be evident in the silence or lack of engagement on the part of stakeholders. At the reporting and dissemination stage, evaluators may be confronted by the most common form of political interference: pressure to change the results of their findings (Morris and Clark 2012; Turner 2003).[7]

Political challenges to evaluation may meld into the domain of ethical challenges — reinforcing the argument above that there is a need to better understand the dynamic interaction between the core domains of evaluation. It underscores the fact that the lines between these domains are not hard and fast. Political questions of power and control quite naturally overlap with ethical questions of right and wrong. This reminds us to be attentive to the possibility that the framing — and addressing — of an issue within one domain may hinder our ability to appreciate its salience within other domains. Jayawickrama illustrates the way a methodological decision about the labelling of a stakeholder group may carry ethical and political consequences. In a discussion of the evaluation of projects 'for' groups labelled as 'vulnerable' he writes, '[the] blanket labeling of whole groups as "vulnerable" pushes us from the methodological into the political'.[7] The consequences of this process are highlighted through a series of field-based vignettes in his work with Strecker (2013).

> Such stories illustrate the myopia of evaluators who selectively seek, and instrumentally use, information that suits pre-conceived notions, while ignoring the realities, problems, and needs of the community within which they are working. However, at the end of the day, it is the voice of the evaluator, and that particular representation of the situation that will shape discourse in academic, policy, and practitioner circles. The particular question, within this context, is: What are the implications of, and responses to this kind of evaluation — methodologically, politically, and ethically?[8]

While the example here points to the political consequences of methodological (or, more accurately, epistemological) choices, the sources of the political obstacles confronting evaluators are diverse. Such obstacles may originate from the international, national, or local arenas. Further, in a digital age, geography is no longer a barrier or buffer from the impacts of international political dynamics. One of the authors recalls a meeting attended as a member a World Bank evaluation team in a remote part of Sri Lanka during an especially militarised phase of the conflicts.[9] Within minutes of the conclusion of the meeting, details of the agenda, issues discussed, and the names of those in attendance were posted on a Liberation Tigers of Tamil Eelam (LTTE) website. Thus, the instantaneous flow of information (media, blogs, Facebook) — whether accurate, inaccurate, or malicious — becomes part of the political force field within which evaluators operate.

11

Domain 3: Conflict Context and Methodology

Methodologies for evaluation in conflict zones have tended to be a continuation of the largely linear approaches to evaluation in non-conflict zones. They are often determined, and driven, by external (usually donor) interests and the need for financial accountability. This orientation is not surprising given the origins of evaluation as a field of practice. As Quinn Patton notes:

> programme evaluation as a distinct field of professional practice was born of two lessons … first, the realization that there is not enough money to do all the things that need doing; and second, even if there were enough money, it takes more than money to solve complex human and social problems. *As not everything can be done, there must be a basis for deciding which things are worth doing.* Enter evaluation. (1997, 11)

Consequently, evaluation becomes a central tool for funders seeking to justify past expenditure, and allocate future resources. The same accountability logic dominates governmental and non-governmental evaluation of initiatives in conflict zones.

In general terms, the (largely financial) accountability imperative of evaluation has tended to overshadow learning functions. Further, standard linear approaches to evaluation (typified by the adherence to logical frameworks that have not been updated since project inception) tend to place more emphasis on tangible, short-term outputs of programme activities than on the subtle — less easily measured — outcomes and impacts of interventions, be they development, humanitarian, or peacebuilding in intention. Obviously, accountability is important; however, the disproportionate focus by funders on accountability may stymie organisational learning — something that is much needed in conflict environments. By framing the conversation around accountability, individuals and groups working for social change in conflict zones are less likely to take risks or to innovate as needed in order to navigate conflict complexities (Duggan 2011, 215).

Standard linear approaches to evaluation … tend to place more emphasis on tangible, short-term outputs of programme activities than on the subtle — less easily measured — outcomes and impacts of interventions.

While efforts are underway to overcome this problem, we are a long way from an integrated solution. The discussion below highlights some of the most important ways that conflict context affects the methodological underpinnings of evaluation under such conditions.

Null hypothesis

One of the most important ways that conflict context affects evaluation methodology is also the least appreciated — or the least publically acknowledged. Under 'normal' conditions, the null hypothesis of an evaluation holds that an intervention has failed to have an impact on the intended outcome. A 'null finding' and 'project failure' are treated synonymously. That is, an initiative (whether a peacebuilding project or a water and sanitation project) is assessed on a continuum of success or failure — where 'failure' is framed as a 'null finding': it had no impact on the desired outcome.

In conflict zones, environmental conditions limit the possibility that a project will have absolutely no broader impact. Project failure in conflict zones is unlikely to be of no impact. There is an increased risk of extra-project impact that will be destructive, and sometimes fatal, i.e., the impact may well be *increased* insecurity, volatility, injustice, and violence. Failure is the education project in northern Uganda that resulted in the kidnapping of students from a newly opened school; it is the increase in the vulnerability of children of sex workers participating in child rights projects in South Asia (Zaveri 2013); and it is the legitimation of gun-based governance structures through Western interventions — as when rebels were used to provide security for humanitarian convoys leaving Mogadishu in the early 1990s (or when Western powers prop up human rights-abusing regimes).

> **Project failure in conflict zones is unlikely to be of no impact. There is an increased risk of extra-project impact that will be destructive, and sometimes fatal, i.e., the impact may well be *increased* insecurity, volatility, injustice, and violence.**

The methodological implication for evaluations in conflict zones is clear: the need to ensure that the scope and tools of evaluation are able — indeed *required* — to probe, explore, and measure peace and conflict impacts beyond short-term, measurable outputs of an intervention. Otherwise, we risk blinding ourselves from seeing those projects where the operation is extolled as a shining success but in which the patient dies, off camera. Or, equally problematic, we evaluate the failure of a project to be the result of the conflict environment, rather than being the result of the exacerbating impact of the project on the conflict conditions. Thus, the evaluation of the case above, where children were kidnapped from a school in northern Uganda, cited failure to be the result of the conflict, rather than a function of the NGO's blueprint approach to education projects which inadvertently gave rise to the opportunity for abductions.[10]

Working without baselines

How do you determine the degree to which change has occurred when there is no initial point of comparison? The problem of scarce or non-existent baseline data plagues most development evaluations. Even where data are available, their quality may be dubious. In conflict zones this problem is heightened. Surveys may not have been undertaken on participants before or during the project for reasons of security, political obstacles, technical constraints, or cost. For example, there may be security risks to the collection of the data. Or distrust may inhibit the sharing of personal information that might be useful to armed actors; or cause participants to misrepresent their actual conditions (e.g., in the belief that that more resources will come to the community, or to downplay the influence of armed actors in their lives). Data have often been destroyed or been simply inaccessible or unreliable for a host of conflict-related reasons, which may include insecurity, censorship, or political competition for control of the process (and findings). And there are almost always limits on the comparability of data within, and across, cases because of differences in protocols and practices in data collection — which raise problems of the generalisability or external validity of data. This is true whether an evaluation is undertaken in conditions of militarised violence (Palestine or Afghanistan), social violence (favelas), or criminalised violence (zones under the control of drug gangs in the global North or South).

One of the ways used to address the absence of baseline data is to ask individuals what their lives (or conditions) were like prior to the intervention being evaluated. While this may be better than nothing, it is subject to vagaries of memory and the cognitive biases of

recollection. Further, in environments where there has been displacement of populations, the representations of past conditions gleaned through interviews will completely omit the input from those who were displaced from the area — for example, interviews that take placed in ethnically cleansed geographies.

Finding appropriate counterfactuals

Evaluators find themselves on thin ice, methodologically, when conditions in conflict zones make it impossible or unethical to employ experimental or quasi-experimental designs, or valid pre- and post-testing of project participants. This raises the challenge of developing and employing ways of gauging the impact of an intervention in the absence of a 'pre-intervention' snapshot of the intended phenomenon/condition to be changed/transformed. Bamberger and colleagues' (2006) work on 'real world evaluation' is helpful in suggesting ways in which this particular challenge might be managed *under non-conflict conditions*. While their analysis travels well between cases of evaluation in the global North and Global South, more work is required to translate and apply it to conflict zones.

Empirically, it might be possible to compare project stakeholders with 'equivalent' stakeholders not involved in the project — assuming that the conflict-zone obstacles of distrust, access, and insecurity can be overcome. For example, a project designed to decrease levels of cross-community conflict at a particular interface in Belfast might be compared with other interface locations in the city, using indicators such as incidents of violence and confrontation, including recreational rioting, vandalism, assaults, property damage by missiles and projectiles over the interface walls, perceptions of (in)security and so on. To the extent that the two communities being compared exhibit similar conditions (economic conditions, prevalence of paramilitary activity, levels of unemployment and educational attainment, and so on), any lower level of violence in the project community *may* be attributed to the intervention. For this quasi-experimental approach to work the evaluator must possess an intimate understanding of the two communities — particularly the local-level dynamics of conflict — in order to rule out the possibility that the variation in levels of violence was not the result of non-project-related factors (or another project).

While the relatively manageable levels of risk and insecurity in Belfast may permit the levels of access required for this kind of evaluation approach, it is a different story in more heavily militarised environments, where access to comparison groups is impossible. In such cases, the use of nuanced counterfactual arguments may need to be developed — where the strength of the evaluation rests in the logic and persuasiveness of the counterfactual case being made, rather than in the empirical strength of data. The question underpinning the evaluation is this: given what we know about the social, political, economic, and security conditions in the project area, what changes would we reasonably expect to have occurred in the absence of the intervention of a project? Such a counterfactual is summed up in an exchange that took place on the east coast of Sri Lanka during the wars:

> 'How do you know that the Butterfly Peace Garden [a socially engaged creative arts project with war-affected children] is having any kind of peacebuilding impact at all?'
>   'I don't know about peacebuilding impact. But I can say that not one of the children who has gone through the Garden has voluntarily joined the rebels.'
>   'How do you know that?'

14

'We know that because we are from these communities and we live in these communities.'
'But over 900 kids have gone through the [nine-month] programme.'
'Nine hundred and forty-three.'[11]

In the context of eastern Sri Lanka at the time — where the voluntary and forced recruitment of children by the LTTE was rampant — this was an extraordinary claim. The confirmation of this fact — and details of interventions for the release of children who had been abducted by the LTTE — offers considerable weight to the argument that in the absence of the Butterfly Peace programme a considerable number of the participating children would have ended up fighting for the rebels.

Proxy indicators in conflict zones

When conflict conditions, such as levels of insecurity or the absence of trust, inhibit the collection of direct measures of the impact of an intervention, it is possible to employ proxy indicators. Particularly noteworthy is their growing use in quantitative studies in conflict-affected areas. Some of these initiatives have made innovative use of increasingly accessible technologies — for example, the use of satellite imaging to measure night-light emissions from villages to track the community-wide developmental impact of ransoms paid to Somali pirates (Shortland 2012); the use of remote sensing technology to monitor changes in land cultivation as a measure of citizen welfare in wartime Liberia (Lidlow 2010). Other proxy indicators may be more economistic and immediate, such as the use of the average cost of an assault weapon on the black market (Killicoat 2007) as an indicator of the weaponisation of society. Of course, none of these indicators is sufficient on its own. But each contributes to the empirical contextualisation of the degree to which an intervention being evaluated may have had an impact.

Grappling with fluid conflict systems

It was noted above that development evaluation practice, in general, tends to be dominated by linear, top-down designs and driven by an audit-focused logic. The drawbacks of this approach are carried into evaluation in conflict zones. However, one of the limitations of realising the potential of evaluation in conflict zones is a narrow appreciation of the range of tools and approaches at our disposal.[12]

Some of the most interesting and important work yet to be done in the study of evaluation in conflict zones involves examining and piloting more creative, adaptive, and responsive evaluation approaches. Approaches such as process-tracing methodologies, outcome mapping, and development evaluation are rooted in complex adaptive systems thinking. Collectively, these various approaches, applied in a broad range of disciplines, are anchored in 'complexity theory'.[13] This may be summarised as follows:

> **Some of the most interesting and important work yet to be done in the study of evaluation in conflict zones involves examining and piloting more creative, adaptive, and responsive evaluation approaches.**

[an approach focusing on] how individuals and organisations interact, relate and evolve within a larger social ecosystem. Complexity also explains *why* interventions may have un-anticipated consequences. The intricate inter-relationships of elements within a complex system give rise to multiple chains of dependencies. Change happens in the context of this

intricate intertwining at all scales. We become aware of change only when a different pattern becomes discernible. (Mitleton-Kelly 2007, 2)

Systems thinking and complexity theory have the potential to highlight and incorporate the central role of context in violently divided societies — its fluidity, ambiguity, non-linearity, contingency, multiplicity, simultaneity, and so on (Hawe et al. 2009; Hendrick 2009; PANOS 2009). Recent examples of efforts to use complexity-informed approaches have been undertaken in evaluations of capacity-building, the social dimensions of HIV and AIDS, community development, and research and communication programmes. While the evaluation research community has made some progress on the development of tools and methods to operationalise complexity thinking, questions remain around modes of practice and applicability.[14] It is our hope that that the evaluation research agenda moves in the direction of piloting such approaches to evaluation in conflict zones.

Domain 4: Evaluation Ethics in Conflict Zones

As the space for evaluation contracts, and the domains of evaluation merge into each other (see above), we see that decisions and actions by evaluators in the realms of the political, the logistical, and the methodological inevitably impinge on the ethical. A very broad range of factors in conflict zones subsidise this process, in particular the absence of those normal oversight structures that would typically condition ethical behaviour, such as rule of law, societal structures, institutional norms, and codes of professional conduct.[15] The increased risk to evaluation stakeholders increases the ethical imperative on evaluators to ensure their safety and well-being during *and after* the evaluation.

Looking for ethical guidance

Ethical Issues related to those who commission and/or use evaluations (Turner 2003)

- Managers or funders trying to influence or control evaluation findings, sometimes including pressure on evaluators for positive results (cited repeatedly), sometimes including pressure to provide "dirt" on a programme

- Conflicts between an organisation's needs and those of the client (when working as an internal evaluator)

- Political interference

- Dissemination or suppression of reports

- Requests to use information gathered for one purpose (e.g. programme improvement) for a different purpose (e.g. accountability)

- Unilateral changes to terms of reference midstream or at time of reporting an evaluation and dealing with the implications for quality and relevance of data collected)

- Surfacing issues of incompetence or poor performance among programme staff

Special Ethical Issues related to dealing with different types of evaluation subjects or participants (Turner 2003)

- Working with indigenous people

- Research with children

- Sensitive topics such as sexual victimisation

- Feeding back results to participants

- Informed consent

- Privacy and confidentiality

- Risks of interview subjects disclosing confidential or inappropriate information in interview

Ethical dilemmas are a fact of life for all evaluators, regardless of whether or not they are working in a conflict context. The pervasiveness of ethical challenges is summed up in a statement by a respondent to a survey by the Australasian Evaluation Society (AES): 'ethical dilemmas are pretty much part of the territory in evaluation' (Turner 2003, 38). Based on a survey of its membership, the AES drafted a list of the ethical challenges confronting evaluators under normal conditions (refer to the text box above).

The ethical challenges confronting evaluators under 'normal' conditions are equally — indeed, *more* — prevalent in conflict zones. The difference is that the consequences of ethical shortcomings or miscalculations are much more significant — and potentially lethal. However, one evaluator's ethical issue might be cast as a commissioner's political or methodological problem (Morris 2008). Both evaluators and commissioners often have moral, practical, and political reasons for wanting to frame ethical conflicts as technical or methodological problems. This becomes extreme in conflict-affected settings — places in which a premium is placed on donor visibility and where there is acute pressure to be seen to be 'doing good' and not exacerbating conflict (particularly in peacekeeping and peacebuilding interventions).

> **The ethical challenges confronting evaluators under 'normal' conditions are equally — indeed, *more* — prevalent in conflict zones. The difference is that the consequences of ethical shortcomings or miscalculations are much more significant — and potentially lethal.**

While ethical issues are acutely important within the context of the conduct of evaluations in conflict zones, evaluators find themselves with few avenues for practical guidance, despite, for example, the most recent guidelines produced by the OECD on the conduct of evaluation in 'settings of conflict and fragility', which seek to 'to promote critical reflection [and] to help fill the learning and accountability gap in settings of conflict and fragility by providing direction to those undertaking or commissioning evaluations and helping them better understand the sensitivities and challenges that apply in such contexts' (OECD 2012, 20). While these much anticipated guidelines consolidate and establish the broad parameters of evaluation *of peacebuilding activities* in conflict zones, there is a conspicuous black hole when it comes to appreciating the

conflict-specific ethical challenges: 'ethics' is mentioned only four times in the hundred-page report (OECD 2012, 21; 38; 90).

The invisibility of ethics in evaluation manuals is not unique to the OECD: the *Impact Assessment Guidelines of the European Commission* (EC 2009) refer to ethics only twice (in a summary table); the American Evaluation Association (2004) *Guiding Principles for Evaluators* refers to ethics once in the context of 'professional ethics';[16] and the European Community Humanitarian Office *Manual for the Evaluation of Humanitarian Aid* (ECHO 1999) has no reference to ethics at all. In each case, these hit-and-run references to ethics are undertaken in a wholly hortatory manner. That is, while they exhort evaluators and evaluation commissioners to behave ethically, they provide no concrete direction for how to do so. Not surprisingly, there is no discussion of the conflict-zone-specific nature of ethical challenges confronting evaluators.

This is not to say that there are no ethical guidelines for the field of evaluation. The Canadian Evaluation Society establishes a set of three standards intended to serve as guiding ethical principles: 'competence, integrity and accountability'. Further, 'evaluators are to act with integrity in their relationship with all stakeholders'.[17] The American Evaluation Association Guiding Principles consist of: systematic inquiry; competence; integrity/honesty; respect for people; and responsibility for general and public welfare.[18] These are essential building blocks for evaluation as a field of practice. In the current context, however, it needs to be emphasised that (1) none of these guidelines is conflict-zone 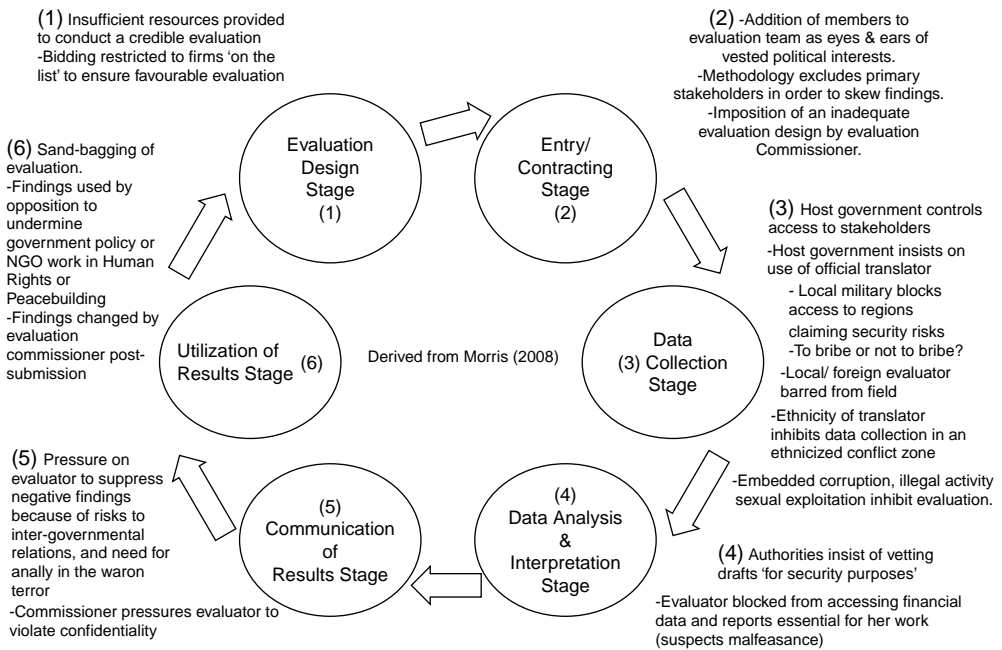specific, and (2) all of them are conspicuously hortatory, leaving the evaluator to their own devices as to how exactly the principles or exhortations should be applied. One of the reasons for this absence can be traced to an inescapable reality: ethics, unlike methodology, do not easily lend themselves to standardisation; no two situations are identical, and we are dealing with human morality and human behaviour.

> **Ethics, unlike methodology, do not easily lend themselves to standardisation; no two situations are identical, and we are dealing with human morality and human behaviour.**

Within the academic literature, it is possible to find some robust discussions of ethical challenges confronted by evaluators in the work of Michael Morris. In a recent review of all the articles published in *Evaluation Practice* and the *American Journal of Evaluation* over the last 25 years, he concludes that there is a 'need for increased empirical research on the ethical dimensions of evaluation, especially as these dimensions are perceived by stakeholders with whom evaluators interact' (Morris 2011, 134). Further, Morris notes that there was, at the time of publication, only one textbook devoted to programme evaluation ethics — Newman & Brown (1996).[19] One very positive development was the establishment, in 2000, of the Ethical Challenges series of the *American Journal of Evaluation*, in which commentators were invited to analyse and discuss the ethical dimensions of a selected case.[20]

It would appear that, although the issue of ethics in evaluation is gradually attracting attention in the literature, there is still much more work to be done to translate this into systematic practice. Such a call for increased ethical sensibilities applies across the field of evaluation.

Morris (2008) helps begin to delineate the different kinds of ethical dilemmas at different stages of the evaluation process in Figure 3. Examples are provided for illustrative purposes.

Figure 3: Examples of Ethical Challenges at Different Stages in the Evaluation Process in Conflict Zones

(1) Insufficient resources provided to conduct a credible evaluation
-Bidding restricted to firms 'on the list' to ensure favourable evaluation

(2) -Addition of members to evaluation team as eyes & ears of vested political interests.
-Methodology excludes primary stakeholders in order to skew findings.
-Imposition of an inadequate evaluation design by evaluation Commissioner.

(6) Sand-bagging of evaluation.
-Findings used by opposition to undermine government policy or NGO work in Human Rights or Peacebuilding
-Findings changed by evaluation commissioner post-submission

(3) Host government controls access to stakeholders
-Host government insists on use of official translator
- Local military blocks access to regions claiming security risks
-To bribe or not to bribe?
-Local/ foreign evaluator barred from field
-Ethnicity of translator inhibits data collection in an ethnicized conflict zone
-Embedded corruption, illegal activity sexual exploitation inhibit evaluation.

(5) Pressure on evaluator to suppress negative findings because of risks to inter-governmental relations, and need for anally in the waron terror
-Commissioner pressures evaluator to violate confidentiality

(4) Authorities insist of vetting drafts 'for security purposes'
-Evaluator blocked from accessing financial data and reports essential for her work (suspects malfeasance)

Evaluation Design Stage (1)

Entry/ Contracting Stage (2)

Utilization of Results Stage (6)

Data Collection Stage (3)

Communication of Results Stage (5)

Data Analysis & Interpretation Stage (4)

Derived from Morris (2008)

For university-based research in conflict zones, projects are largely vetted by ethics review boards of one form or another. While this process is not without its problems — e.g., lack of quality control and absence of monitoring and enforcement capacity — it is nonetheless a mechanism that undertakes a formal ethical assessment of all research proposals involving human subjects. It is a process undertaken by an entity (often university or government based) that technically possesses the authority to request changes to the project, or to reject it all together based on ethical considerations. Yet, such mechanisms do not exist for evaluation interventions — though arguably, in some cases, evaluation advisory groups have exercised an informal ethics review function (VeLure Roholt & Baiserman 2012). Indeed, to the best of our knowledge, they do not exist for other interventions in conflict zones either — such as development, peacebuilding, or humanitarian interventions. Even so, there is a curious grey zone between ethics guidelines for evaluators, evaluation advisory groups, and contractual obligations/ requirements imposed on evaluators by their clients.

The power imbalance between the evaluator and their client has attracted the attention of evaluators for years. As most evaluators know, at the end of the day an unresolved ethical or 'politico-ethical' conflict with a client can have detrimental professional consequences (e.g. the withdrawal of remuneration; reputational damage; and blacklisting from future contracts). Despite the ethical guidelines and principles *for evaluators*, there are no corresponding enforceable standards for evaluation commissioners or funders. Until there are more robust mechanisms for evaluation ethics review, evaluators are left to work it out for themselves — within the constraints of contract law and vague guidelines exhorting them.

19

It soon becomes clear that, in order to be able to act ethically, evaluation actors need:

- an empirically grounded appreciation of the complex, multi-layered dimensions of the context within which they are operating, and an understanding of the potential implications this may have for evaluation;

- A clear sense of their own ethical compass, and the kinds of conflict zone-specific ethical challenges that have arisen in the past — and how they were addressed (or not).

While evaluators may not be able to prevent or avoid every ethical challenge lurking in conflict zones, training and communities of practice are essential to enable us to better anticipate and defuse some of them — and, importantly, to be prepared for the unavoidable ethics crises that may arise. The use of scenario-based learning and 'ethics clinics' for evaluators and for commissioners helps to build up that ethical muscle or skill set. Based on our own teaching, this learning process is particularly effective when peer support and real-life experiences are systematically integrated into the mix.

Training on its own, however, is not the complete answer. It would be helpful, as a start, to translate ethical guidelines from abstract hortatory principles into practical and relevant manuals and tools. Further, evaluation guidelines and policy documents must start taking ethics seriously by including — if not integrating — them *centrally* throughout the entire evaluation process.

Conclusion

If one thing becomes clear from the discussion in this article, it should be this: the evaluation of our interventions in conflict zones is the Rosetta Stone for understanding and systematically strengthening those social, political, and economic substructures needed to support locally determined paths to peaceful, prosperous, and just futures. In the absence of good evaluation, we are left to make decisions based on impressionistic or anecdotal assessments, or worse: opaque political, economic, or particularistic interests. Evaluation is not, however, a silver bullet, any more than development aid or peacebuilding interventions are silver bullets. Evaluation requires time and resources. Poor evaluation is not only a waste of both of these; it can have negative consequences for evaluation actors, both local and international.

Evaluation efforts to identify and understand these links are entangled in a nest of political and economic interests that interact with the conceptual, methodological, ethical, and practical challenges that define this area of enquiry. Until there is a culture of systematic conflict-specific evaluation of interventions in conflict zones (sic), we limit our ability to understand the impact of our interventions — good, bad, or indifferent. This article and the special issue of which it is a part are intended to be a step in the direction of changing this reality. More specific take-away points from this article include the following.

Context matters

The extreme nature of the conflict environment shapes and amplifies the challenges of conducting evaluations and, importantly, the consequences of each decision made in the process. Evaluators *and commissioners* must be able to tease out the inevitable peace,

conflict, or mixed impacts of any and all interventions in conflict zones — this includes the evaluation itself, as well as the initiative being evaluated.

Evaluation is fundamentally a political exercise

The conduct of evaluation in conflict zones is embedded in the political dynamics of the environment. Evaluators are faced with multiple pressures emanating from multiple, intersecting conflicts, as well as power imbalances, donor-driven approaches, and their own value systems. In this context, evaluators should be prepared for political complications throughout the evaluation process.

A broader skill set is required of evaluators in conflict zones

Evaluation in conflict zones requires a skill set that goes beyond the usual social science approaches and tools at the disposal of evaluators. In addition to the usual technical competencies of evaluators, they (or their team) need to possess:

- a well-calibrated moral compass, and an appreciation of the conflict-zone-specific ethical challenges;
- political sensitivities, diplomacy, and conflict resolution skills;
- peace and conflict research skills;
- anthropological, historical, and political sensibilities;
- in militarised zones, a technical knowledge of the structures, strategies, weapons, and behavioural patterns of all armed actors;
- knowledge and appreciation of the intersection of the political and ethnographic at local levels;
- cultural competence and cultural humility.

Methodology

Conventional, linear approaches to evaluation are often insufficient in conflict zones. The introduction and growing practice of creative, flexible, and adaptive evaluation approaches rooted in systems and complexity thinking would help generate robust, useful findings. In this process, meta-evaluation (that is, the evaluation of evaluations) would be helpful in rebuilding and reshaping evaluation standards and practice in conflict zones by fusing theory, methodology, and practice.

Extreme ethics

Extreme context is infused with extreme ethical implications — more risks, greater risks, and greater consequences of all decisions and actions. Each stage of the evaluation process should be monitored very closely through a politico-ethical lens. Much work remains to be done in examining ethical challenges in conflict zones and in finding strategies to anticipate or address them.

Ethical frameworks, guidelines, and standards should continue to be systematically and periodically assessed. Where appropriate they should be reworked to reflect local values (for example, the African Evaluation Association Guidelines). These ethical frameworks are part of an evolving process of self-examination by the community of evaluators

21

within a global multicultural context, and should be revisited prior to each new evaluation to ensure they are calibrated with a critical local lens.

KENNETH BUSH is a Research Affiliate with the National Centre for Peace and Conflict Studies, Otago University.

COLLEEN DUGGAN is Senior Program Analyst, Corporate Strategy and Evaluation Division, International Development Research Centre (IDRC), Ottawa.

Endnotes

[1] Noteworthy recent initiatives include the establishment of the Learning Portal for Design, Monitoring and Evaluation for Peacebuilding (http://dmeforpeace.org/) and the US Institute for Peace funded Peacebuilding Evaluation Project (http://www.allianceforpeacebuilding.org/?page = workpep). A particularly important contribution to this field of work is the OECD's (2012) *Evaluating Peacebuilding Activities in Settings of Conflict and Fragility.*

[2] The term and idea of PCIA is introduced in Bush (1998). Elaboration is provided in Bush (2003; 2005). For a series of critical debates on the theory and practice of PCIA, see the two sets of commissioned essays in the *Berghof Centre Handbook for Constructive Conflict Transformation*: http://www.berghof-handbook.net/dialogue-series/

[3] For example, the research undertaken, or supported, by International Alert and IDRC — see International Alert et al. (2004). This PCIA research has also been incorporated into programmes such as the PEACE III Programme in Northern Ireland. See: http://www.seupb.eu/Libraries/PEACE_III_Practical_Project_Guidlines/PIII_paper_practical_project_guidelines_090519__Aid_for_Peace_Approach.sflb.ashx

[4] Our use of the term 'impact' corresponds with the commonly accepted definition: 'Positive and negative, primary and secondary long-term effects produced by [an] … intervention, directly or indirectly, intended or unintended' (OECD 2002).

[5] In development studies, see Scheyvens and Storey (2003) and Devereux and Hoddinott (1993). In the anthropology of violence, see Kovats-Bernat (2002) and Belousov et al. (2007). More practically, see UN (1998) and Cutts and Dingle (1995).

[6] This is one of many stories emerging from our ongoing project on the 'ethical tipping points of evaluators in conflict zones'. Evaluators spoke on the condition of anonymity. Skype Interview Date, 14 April 2013.

[7] This passage is culled from the unpublished research prospectus prepared by Janaka Jayawickrama for a project on the evaluation of research in conflict zones (see Jayawickrama & Strecker, 2013).

[8] This passage is culled from the unpublished research prospectus prepared by Janaka Jayawickrama for a project on the evaluation of research in conflict zones (see Bush & Duggan forthcoming).

[9] Kenneth Bush, field notes, 2002.

[10] The heart-breaking footnote to this case is that the parents of the village subsequently burned the school to the ground. As far as I know, the large NGO responsible for the project is still receiving funding from bilateral agencies for projects using the same flawed blueprint.

[11] Kenneth Bush interview with Father Paul Satkunayagam, Director and Co-founder of the Butterfly Peace Garden, February 2002.

[12] For excellent overviews, discussion, and tools of the panoply of approaches see Rick Davies, *Monitoring and Evaluation News* (http://mande.co.uk/) and *The Learning Portal for Design, Monitoring, and Peacebuilding* (http://dmeforpeace.org/).

[13] Ibid.

[14] The IDRC's outcome mapping is one example of a methodology that uses a systems approach to untangle the problems of evaluating research. Applying complexity theory has been more challenging mainly in terms of translating theory into a useable framework for practitioners. See Ramalingam et al. (2008) and Verkoren (2008).

[15] This is *not* arguing (or supporting the argument) that militarised conflict zones are anarchic or lacking in social, political, or economic structures. Rather, we are arguing that such structures may be subordinated to, or transformed by, protracted dirty war.

[16] D-2: 'Abide by current professional ethics, standards, and regulations regarding confidentiality, informed consent, and potential risks or harms to participants'.

[17] For a copy of the standards see: http://www.evaluationcanada.ca/site.cgi?en:6:10

[18] http://www.eval.org/publications/guidingprinciples.asp

[19] Also noteworthy in this context is Chapter 11, on ethics, in Church and Rogers (2006).

[20] For details, see: http://www.eval.org/publications/AJEcontribcats.asp

References

American Evaluation Association 2004, *AEA Guiding Principles for Evaluation*: http://www.eval. org/publications/guidingprinciples.asp

Bamberger, M., **Rugh, J.** & **Mabry, L.** 2006, *Real World Evaluation: Working under Budget, Time, Data, and Political Constraints*, Thousand Oaks, CA: Sage.

Belousov, K. et al. 2007, 'Any Port in a Storm: Fieldwork Difficulties in Dangerous and Crisis-Ridden Settings' in *Qualitative Research* 7: 2: 155–175.

Bush, K. 1998, *A Measure of Peace: Peace and Conflict Impact Assessment (PCIA) of Development Projects in Conflict Zones*, Ottawa: International Development Research Centre.

Bush, K. 2003, 'PCIA Five Years on: The Commodification of an Idea' in *The Berghof Handbook for Conflict Transformation*: http://www.berghof-handbook.net/documents/publications/dia logue1_bush.pdf

Bush, K. 2005, 'Field Notes: Fighting Commodification and Disempowerment in the Development Industry: Things I Learned About Peace and Conflict Impact Assessment (PCIA) in Habarana and Mindanao' in *The Berghof Handbook for Conflict Transformation*: http://www.berghof-ha ndbook.net/articles/PCIA_addKB.pdf

Bush, K. & **Duggan, C.** forthcoming, *Evaluation in Extremis: The Politics and Impact of Research in Violently Divided Societies*, Delhi: Sage.

Church, C. & **Rogers, M.** 2006, 'Chapter 11: Ethics' in *Designing for Results: Integrating Monitoring & Evaluation in Conflict Transformation Programs*, Washington: Search for Common Ground.

Cutts, M. & **Dingle, A.** 1995, *Safety First — Protecting NGO Employees Who Work in Areas of Conflict*, London: Save the Children.

Devereux, S. & **Hoddinott, J.** 1993 eds, *Field Work in Development Countries*, Boulder CO: Lynne Rienner Press.

Duggan, C. 2011, 'Aligning for Impact: Human Rights' in *Journal of Human Rights Practice* 3: 2: 214–219.

EC 2009, *Impact Assessment Guidelines of the European Commission*, SEC(2009) 92.

ECHO 1999, *Manual for the Evaluation of Humanitarian Aid*: http://www.cihc.org/members/ resource_library_pdfs/8_Programming/8_2_Evaluation/echo-manual_for_he_Evaluation_of_ Humanitarian_Aid-1999.pdf

Fitzpatrick, J.L. & **Morris, M. eds**. 1999, *Current and Emerging Ethical Challenges in Evaluation: New Directions for Evaluation, No. 82*, San Francisco, CA: Jossey-Bass.

Hawe, P., **Bond, L.** & **Butler, H.** 2009, 'Knowledge Theories Can Inform Evaluation Practice: What Can a Complexity Lens Add?' in *New Directions in Evaluation* 124: 89–100.

Hendrick, D. 2009, 'Complexity Theory and Conflict Transformation — an Exploration of Potential and Implications', Centre for Conflict Resolution Department of Peace Studies Working Paper 17, University of Bradford: http://www.brad.ac.uk/acad/confres/papers/pdfs/CCR17.pdf

International Alert et al. 2004, *Conflict-Sensitive Approaches to Development, Humanitarian Assistance and Peacebuilding: A Resource Pack*, London: International Alert.

Jayawickrama, J. & **Strecker, J.** 2013, 'The Ethics of Evaluating Research. Views from the Field' in Bush, K. & Duggan, C. eds, *Evaluation in Extremis: The Impact and Politics of Research in Violently Divided Societies*, Delhi: Sage.

Killicoat, P. 2007, 'Weaponomics: The Global Market for Assault Rifles', Post-Conflict Transitions Working Paper No. 10, World Bank Policy Research Working Paper, WPS 4202: http://www-wds.worldbank.org/external/default/WDSContentServer/IW3P/IB/2007/04/13/000016406_20070413145045/Rendered/PDF/wps4202.pdf

Kovats-Bernat, J.C. 2002, 'Negotiating Dangerous Fields: Pragmatic Strategies for Fieldwork Amid Violence and Terror' in *American Anthropologist, New Series* 104: 1: 208–222.

Lidow, N. 2010, 'Rebel Governance and Civilian Abuse: Comparing Liberia Rebels Using Satellite Data'. paper prepared for the American Political Science Association (APSA) Conference, Washington, September.

Mitleton-Kelly, E. 2007, 'The Emergence of Final Cause' in Aaltonen, M. eds, *The Third Lens: Multi-ontology Sense-Making and Strategic Decision-Making*, Aldershot: Ashgate.

Morris, M. 2008, 'Ethics and Evaluation' in Morris, M. eds, *Evaluation Ethics for Best Practice. Cases and Commentaries*, New York: Guilford Press.

Morris, M. 2011, 'The Good, the Bad, and the Evaluator: 25 Years of AJE Ethics' in *American Journal of Evaluation* 32: 1: 134–151.

Morris, M. & **Clark, B.** 2012, 'You Want Me to Do What? Evaluators and the Pressure to Misrepresent Findings' in *American Journal of Evaluation* 34: 1: 57–70.

Newman, D.L. & **Brown, R.D.** 1996, *Applied Ethics for Programme Evaluation*, Thousand Oaks, CA: Sage.

OECD 2002, *Glossary of Key Terms in Evaluation and Results Based Management*, http://www.oecd.org/development/peer-reviews/2754804.pdf, accessed 22 June 2013.

OECD 2007, 'Encouraging Effective Evaluation of Conflict Prevention and Peacebuilding Activities: Towards DAC Guidance' in *OECD Journal on Development* 8: 3, doi: 10.1787/1996580x

OECD 2008, *Guidance on Evaluating Conflict Prevention and Peace Building Activities*, Paris.

OECD 2010, *Quality Standards for Development Evaluation*, Paris: OECD, http://www.oecd.org/dataoecd/55/0/44798177.pdf, accessed 19 December 2011.

OECD 2011, *Evaluating Conflict Prevention and Peacebuilding Activities. Lessons from the Application Phase. Chairman's Summary*, Paris: OECD, http://www.oecd.org/dac/evaluationofdevelopmentprogrammes/dcdndep/48092571.pdf

OECD 2012, *Evaluating Peacebuilding Activities in Settings of Conflict and Fragility: Improving Learning for Results*, Paris: OECD, http://dx.doi.org/10.1787/9789264106802-en

PANOS 2009, 'How Can Complexity Theory Contribute to More Effective Development and Aid Evaluation', dialogue at the Diana, Princess of Wales Memorial Fund, London, 3 July: http://evaluationrevisited.files.wordpress.com/2010/05/panos-london-complexity-and-evaluation-dialogue.pdf

Patton, M.Q. 1997, *Utilization Focused Evaluation: The New Century Text*, London: Sage, 3rd edn.

Ramalingam, B., **Jones, H.**, **with Reba, T.** & **Young, J.** 2008, *Exploring the Science of Complexity: Ideas and Implications for Development and Humanitarian Efforts*, London: Overseas Development Institute: http://www.odi.org.uk/sites/odi.org.uk/files/odi-assets/publications-opinion-files/833.pdf

Scheper-Hughes, N. & **Bourgois, P.** 2004, 'Introduction' in Scheper-Hughes, N. & Bourgois, P. eds, *Violence in War and Peace: An Anthology*, Oxford: Blackwell.

Scheyvens, R. & **Storey, D.** 2003 eds, *Development Field Work. Practical Guide*, Delhi: Sage.

Shortland, A. 2012, *Treasure Mapped: Using Satellite Imagery to Track the Developmental Effects of Somali Piracy*, London: Chatham House: http://www.chathamhouse.org/sites/default/files/public/Research/Africa/0112pp_shortland.pdf

Turner, D. 2003, 'Evaluation Ethics and Quality: Results of a Survey of Australasian Evaluation Society Members', AES Ethics Committee, November: http://www.aes.asn.au/images/stories/files/About/Documents%20-%20ongoing/ethics_survey_summary.pdf

UN 1998, *Security in the Field*, New York: Office of the UN Security Coordinator.

VeLure Roholt, R. & **Baiserman, M.** 2012, 'Special issue on Evaluation Advisory Groups' in *New Directions in Evaluation* 136.

Verkoren, W. 2008, 'Debating Complexity: The Value of Complexity for Development' in *The Broker — Connecting Worlds of Knowledge*: http://www.thebrokeronline.eu/en/Articles/Debating-complexity

Zaveri, S. 2013, 'Evaluation and Vulnerable Groups: Forgotten Spaces' in Bush, K. & Duggan, C. eds, *Evaluation in Extremis: The Impact and Politics of Research in Violently Divided Societies*, Delhi: Sage.

25

'If they can't do any good, they shouldn't come': northern evaluators in southern realities

Janaka Jayawickrama
Published online: 19 Sep 2013.

PLEASE SCROLL DOWN FOR ARTICLE

'IF THEY CAN'T DO ANY GOOD, THEY SHOULDN'T COME': NORTHERN EVALUATORS IN SOUTHERN REALITIES

JANAKA JAYAWICKRAMA

Abstract

Based on research and evaluation experiences from Sri Lanka, Malawi, Sudan and Pakistan, this paper reflects on critical implications of ethics at local level. Providing various examples, the paper invites the reader to think through solutions to their own questions and challenges in the realm of evaluation ethics. Further, this paper examines the ethics of evaluation through the lens of the author's own experiences as a researcher and evaluator, which is shaped by his particular positionality and by what he has learnt working in the space within and between the European university system and violently divided communities. From this standpoint, this paper has managed to distinguish some of the central challenges to evaluators working in violently divided communities.

Introduction

There is a growing recognition that the humanitarian aid sector as well as the wider international aid system will have to adapt to deal with fast change, complexity, uncertainty and danger — a more than challenging environment for policymakers and implementing agencies. With this in mind, those with evaluation roles and responsibilities will also have to make related adjustments to their moral and ethical frameworks when it comes to disaster and conflict-affected communities. Listening to affected communities will be ever more essential, but, as ethical concepts change, ever more difficult.

There are several ethical issues that are common across all types of research, including evaluation, which in itself is a form of research. Ethical dimensions may touch upon research design, methodology, sources of funding, and methods in reporting data. For the evaluator, ethical considerations are present throughout the evaluation process, including during: entry/contracting, design of the evaluation, data collection, analysis and interpretation, communication of results, and the utilisation of results (Morris 2007). Importantly, although development and humanitarian agencies are increasingly

---

*This paper draws some of its reflections from research undertaken by J. Jayawickrama, J and J. Strecker in the context of a larger research project with INCORE and IDRC, between 2012 and 2013.

26

supporting evaluation research, formal, binding, systemic, ethics review mechanisms are not in place — although non-binding codes of conduct, with no enforcement mechanisms, have been developed.

Research ethics are based on the underlying principles of autonomy, beneficence and justice (Orb et al. 2000). These key ethical principles are designed to protect and respect research participants, doing good for others and preventing harm and promoting fairness (Capron 1989; Raudonis 1992). Within the context of evaluation ethics, Morris adds fidelity and non-maleficence (Morris 2007), explaining that evaluators are expected to maintain fidelity by acting in good faith and ensuring they are loyal, honest and keep their promises (Newman & Brown 1996). Non-maleficence, or the 'Do no harm' principle, exhorts evaluators to avoid inflicting injury on others (either physical or psychological), and to 'protect individuals from exposure to the risk of harm' (Morris 2007, 5). For those situations where harm is unavoidable, the evaluator is expected to manage and reduce harm (where possible) and should maintain a reasonable expectation that the harm incurred will be compensated by the benefits of the evaluation (Morris 2007).

The aim of this article is to discuss critical implications of ethics at local level as well as positive examples from evaluations in violently divided societies to provide an implicit platform for readers to think about solutions to their own questions and challenges on evaluation ethics. I examine the ethics of evaluation in a violently divided society context, for the benefit of evaluators and researchers, through the lens of my own experiences as a researcher and evaluator. This is shaped by my particular positionality and by what I have learnt working in the space within and between the European university system and violently divided communities. From this particular vantage point, I am able to discern some of the central challenges to evaluators working in violently divided communities; principally the challenge of how to transform evaluation guidelines, frameworks and practices in order to recognise affected communities as equal partners of change.

Developed World's Evaluation Principles in Complex Emergencies

Within the Western academic and political traditions (two traditions that are intimately connected) ethics and ethical discourse are derived mainly from scientific knowledge systems that have discounted and marginalised non-Western systems of knowing.[1] At the same time this scientific ethics and ethical discourse often pillage indigenous knowledge systems through various forms of intellectual and geopolitical colonisation and imperialism. Whether we use the language of infantilisation or of orientalisation of the Other the process is the same.[2] This division of knowledge and Western ethics is often intended to bring good to violently divided developing countries. This does not necessarily involve a simplistic notion of North vs. South. There are many Northern researchers with respect for and understanding of Southern realities, and there are many Southern researchers who have been trained in Northern universities and have little or no respect towards Southern realities.

In the field of evaluation, evaluators are not guided by ethical review boards and processes, but are still required to abide by an array of professional principles and frameworks that are articulated and updated by regional and international evaluation associations. Some of the most prominent frameworks include: American Evaluation Association (AEA) Guiding Principles for Evaluators (2004), Australasian Evaluation Society (AES) Guidelines for the Ethical Conduct of Evaluations (2012), Canadian

Evaluation Society Guidelines for Ethical Conduct (2012), African Evaluation Association (AfrEA) African Evaluation Guidelines (2002), French Evaluation Society Charter of Evaluation Guiding Principles for Public Policies and Programmes (2003), Organisation for Economic Cooperation and Development Assistance Committee (OECD-DAC) Quality Standards for Development Evaluation (2010) and United Nation Evaluation Group (UNEG) UNEG Ethical Guidelines (2008).

Many of these frameworks draw from one another, and all of them aim to promote ethical practice and improve evaluation theory and use, by providing guidance and awareness pertaining to ethical issues prominent in evaluation. These frameworks have been proven to provide useful warning of ethical dilemmas in evaluation, but they do not provide a 'blueprint' of how to approach and respond to particular situations (Morris 2007). Several of the frameworks have also been criticised for having a degree of ambiguity. One review of the AEA's Guiding Principles suggested, 'the Principles in particular seem so open to interpretation that a wide range of values, preferences, and opinions can be projected onto them' (Datta 2002, 195 — cited in Morris 2007). While this critique is valid, it is this same openness that enables these principles to be transferred from one context to another. The usefulness of these ethical frameworks is therefore variable, depending on appropriate interpretation and implementation by the evaluator. If these principles are not applied through an appropriate mechanism, they pose the same risk as ethics boards: imposing externally generated principles onto local participants and projects, which may actually cause harm, by subordinating local needs and realities or by creating Southern subservience to Northern ethical principles and agendas.

Within violently divided contexts, the concept of codifying a strict set of unified ethical principles is simply unrealistic and undesirable. Within violently divided contexts, the concept of codifying a strict set of unified ethical principles is simply unrealistic and undesirable.This is because each situation in violently divided societies provides unique social, political and cultural challenges and thus finding and applying a uniform ethical framework is extremely challenging. There are two reasons for this; one is the social, political and cultural differences between the violently divided context and the evaluator. The second relates to the timing requirements of the organisation that commissioned the evaluation. For example, using OECD-DAC's evaluation criteria (2010) — impact, efficiency, effectiveness, relevance and sustainability — the evaluator faces a two-pronged challenge. To evaluate a project through these criteria requires time; time to learn the local social,

Within violently divided contexts, the concept of codifying a strict set of unified ethical principles is simply unrealistic and undesirable.

political and cultural situation, as well as time to cultivate engagement with the community or project beneficiaries. In most cases this is not possible, due to tight timescales or budget limitations, constraints that plague most evaluations. The evaluator walks in and out of the project community; at best, he or she can only hope to gather good-quality data; however, the quality, validity or utility of these data may be questionable if the necessary relationships of trust do not exist between the evaluator and the evaluation's 'subject' or stakeholder.

In brief, the time constraints in Western evaluation, frequently imposed by budget constraints, mean that it is virtually impossible for external evaluation to make judgements on the relevance, sustainability and especially the impact of any programme intervention.

Northern vs. Southern: Critiques and Recommendations

> The Governor of She told Confucius, 'Among my people, there is a man of unbending integrity: when his father stole a sheep, he denounced him.' To this Confucius replied, 'Among my people, men of integrity do things differently: a father covers up for his son, a son covers up for his father — and there is integrity in what they do.' (Quoted in Sen 1999, 235)

Along the lines of what Confucius suggested, this article is examining the different views of ethics in the North and South. However, as explained in the above discussion, Northern ethical frameworks overlook the ethics in Southern practices. There are obvious differences (accepting that there are similarities too) in the underlying values of ethics on both sides and it is important to examine these differences in order to develop better frameworks.

The following sections critique current ethical practices and begin to suggest ways of resolving them. Concepts such as vulnerability and 'Do not harm' are gospel in the current Western ethical frameworks, yet often overlook local ethical values, frameworks and practices and consciously or unconsciously support the evangelising of local practitioners. While the sections on vulnerabilities and 'Do no harm' critique current ethical policy and practice, the sections on local lenses and ethical frameworks offer an alternative direction for evaluators and researchers to think about their own solutions to the challenges in the field.

Navigating vulnerabilities

Conducting evaluations in violently divided societies often means that evaluators will be working with 'vulnerable' stakeholder groups. Vulnerability as an organising concept is constructed differently by different donor agencies and evaluation associations — what might be referred to as 'the politics of vulnerability'. Vulnerability as an organising concept is constructed differently by different donor agencies and evaluation associations — what might be referred to as 'the politics of vulnerability'.Unfortunately, evaluation frameworks provide little guidance on how to approach vulnerability. UNEG's ethical principles only vaguely reference vulnerable groups, noting that all evaluations must comply with legal codes:

> Compliance with codes for vulnerable groups. Where the evaluation involves the participation of members of vulnerable groups, evaluators must be aware of and comply with legal codes (whether international or national) governing, for example, interviewing children and young people. (UNEG 2008, 7)

While it is important for codes and protocols to be followed, these guidelines provide little direction for evaluators, and focus attention on the pathology of vulnerability rather than the reasons for vulnerabilities in violently divided societies, such as inequality. As a result, many evaluators who have a background in research tend to fall back on research ethics guidelines, since these are what they know best. Evaluation is, however, distinct from research in several different ways, the most pertinent being that evaluation drives immediate decision-making and often has immediate consequences for the people and organisations on the ground.

29

Evaluations that only deal with the legalities of working with vulnerable stakeholders, and do not adopt a local lens, tend to reinforce inequalities by focusing on the vulnerability of the participant rather than acknowledging the individual's knowledge and agency.

> My dear son, we may be poor, we may be illiterate and living in difficult conditions, but we are not stupid and not a bunch of idiots.[3]

The quote above was from one of my first experiences with a community member who replied to my questions about vulnerability. This harsh reply made me realise that certain terms are embedded with attitudes and approaches that pathologise and incapacitate communities. I realised that, by concentrating on negativities and vulnerabilities in people's lives, I was attempting to cast them as weak and broken rather than strong and capable of dealing with uncertainty. How ethical is it to label these communities as vulnerable, when they are struggling effectively to maintain everyday life? As Kleinman (2006) argues, they may look vulnerable and fragile from an outside point of view, but in the midst of the worst horrors they indeed continue to live, to celebrate and to enjoy.

While these questions remain unresolved, there is increasing movement away from victimisation and vulnerability pathologies, and towards recognition of the structures that manifest them. The AES Guidelines for Ethical Conduct, for example, have rightly moved beyond the vulnerability rhetoric by discussing *inequalities* as opposed to vulnerabilities.

> Account should be taken of the potential effects of differences and inequalities in society related to race, age, gender, sexual orientation, physical or intellectual ability, religion, socio- economic or ethnic background in the design conduct and reporting of evaluations. Particular regard should be given to any rights, protocols, treaties or legal guidelines which apply. (AES 2012, 9)

Unfortunately, despite the fact that this ethical framework transcends vulnerability pathology by acknowledging the inequalities that underlie it, there is still little direction provided beyond deferring to secondary texts and protocols.

According to Ford et al. (2009), evaluations conducted in violently divided societies have a higher potential for exploiting situations with different power balances in ways that could end up compromising or denying rights of people. This is particularly true in situations where the programme being evaluated is tied to the delivery of aid or life-saving services. In these situations, the power differential is a significant factor, since participants often know or presume the potential consequences for future funding and their livelihood (Duggan 2012). Given this context many individuals feel obligated to participate even if their participation would serve to disadvantage them. For example, a recent study in Darfur uninten-

Failure to assess appropriate timing and methods for evaluations provides another example where outsiders' limited contextual knowledge leads to unethical situations that reinforce inherent power differentials and inequalities.

tionally scheduled interviews during the times of food distribution, which placed participants in an unfair position of having to choose between one or the other (Ford et al. 2009). Participants may also be placed in compromising situations where they feel obligated to answer questions, which could lead to increased distress or the reliving of

traumatic events. Failure to assess appropriate timing and methods for evaluations provides another example where outsiders' limited contextual knowledge leads to unethical situations that reinforce inherent power differentials and inequalities. Failure to assess appropriate timing and methods for evaluations provides another example where outsiders' limited contextual knowledge leads to unethical situations that reinforce inherent power differentials and inequalities.

These examples speak to the importance of applying a critical local lens when designing and conducting evaluations. AES's Guidelines draws awareness to issues of inequality and moves away from the oppressive rhetoric of vulnerability. However, as noted, they fail to provide evaluators with further direction of how to approach these potential ethical dilemmas. Scholars such as Skerry have stressed the importance of thoughtfulness and creativity when involving 'vulnerable' populations in studies (Skerry 2000 quoted in Phillips & Morrow 2005, 65). However, once again these suggestions fail to take account of the insider–outsider dilemma. In external evaluation, the evaluator, as the outsider, is often positioned in a place of privilege with respect to his or her participant. It is therefore vital that evaluations are conducted in a way that is not extractive but empowers the voice of participants, by sincerely acknowledging their agency and knowledge as insiders. In order to do this, evaluations should be designed with a critical local lens, which should take into account the layers of inequalities, and the environmental structures, which reinforce the power dynamics that the evaluator has been dropped into. This is by no means an easy task. Evaluators are often operating under severe time and resources constraints, working to a timeline set by a client in a distant capital. This places limitations on their ability to design and engage in the processes often needed to put in place respectful relationships with evaluation subjects.

Do no harm: 'If they can't do any good, they shouldn't come'

> No one within our community requested these international organisations to come and help us. We have been surviving the conflict since the 1980s and disasters since the 1950s. Before 1990, we were helping each other and the few organisations in our area were listening to us. Now, it is different — all these foreigners and their assistant Sri Lankans who come in Land Cruisers with questionnaires only want our information. Then they disappear and a new group comes. I think that if they can't do any good, they shouldn't come.[4]

In the mid-1990s, the concept of 'Do no harm' became the motto of humanitarian policy and practice. Although the concept has been part of the medical field's Hippocratic Oath since the late 5th century BC, it entered the humanitarian lexicon through the work of James Orbinski, and was adapted and globally promoted by Mary B. Anderson and her Collaborative for Development Action (CDA).

The Do No Harm Project began in 1993, with the aim of recognising ways to deliver humanitarian and/or development assistance in conflict-affected communities. The driving concept behind this idea is that when frontline workers understand the patterns of harmful assistance they can create opportunities to overcome the conflict by reducing harmful practices and increasing positive effects. In this way, they can achieve their mandates to assist, avoid doing the harm that has been done in the past, and add the influence of their presence and assistance to the forces within societies that reconnect people rather than separate them (CDA 2007).

31

While this is the global mandate that distinguished the Do No Harm concept, at times the reality on the ground has been markedly different. During my field experience with communities in Sri Lanka, Malawi, Sudan, Jordan, Darfur and Pakistan, community members provided very different narratives from those of research and evaluation outsiders:

> A team of researchers came to our camp and wanted to gather our experience with the war. They wanted to know our losses of loved ones and properties. By that time we had enough with these researchers who just talk to us and go away. But the Agency insisted us to talk to these researchers too. So, we agreed. These researchers were very difficult — they were not interested about our current problems in the camp, but wanted us to tell all that aspects of our experiences where we felt so sad to remember them. They were very pushy to get all what they want. At the end there were crying women and upset and angry men.[5]

Such stories illustrate the myopia of researchers who selectively seek, and instrumentally use, information that suits preconceived notions, while ignoring the realities, problems and needs of the community within which they are working. The voices from this community suggested that researchers and evaluators who come to collect information from them should actually listen to them and address their issues rather than just focus on pleasing their donors and accomplishing their agendas.

> As an academic, I know how difficult it is for these researchers to understand the complicated situations the communities are in, while formulating their research agendas. They have to make their funders happy. Then they have to follow their ethical frameworks and research objectives from their institutions. However, what they should think about is that we are also living human beings. These are our lives and making dishonest judgements about our situations is unethical and immoral. I don't know how they sleep at night. Conducting social research is not just a job, but a responsibility towards the research participants. With all the good intentions, you can still damage us.[6]

It is vital that evaluators acknowledge that they are not assessing rocks and soil; they are engaging with human beings who have experienced conflict or disaster and may not wish to relive these experiences through interviews. Participants of an evaluation trust the evaluator to acknowledge or share their experiences and future aspirations. Although, one could argue that what is collected is simply field data from the evaluation subjects, many communities recognise a different relationship. For them, the moment these communities share their stories the evaluator becomes part of them. This establishes an unwritten agreement that the researcher will respect and do justice to these stories. Evaluators may not meet these participants again, but their responsibility towards their stories remains forever.

It is vital that evaluators acknowledge that they are not assessing rocks and soil; they are engaging with human beings who have experienced conflict or disaster and may not wish to relive these experiences through interviews.

The job of the evaluator is to uncover the contributions or strength of the project being assessed. Evaluators are therefore faced with a difficult task of assessing whether the responsibilities of the researchers or development workers have been adequately met,

while at the same time balancing their own responsibility to the evaluation's stakeholders. This delicate balance is part of the reason why the 'Do no harm' mantra has been adopted into evaluations' ethical frameworks. Evaluations, to remain ethical, must protect participants from unnecessary exposure to harm (Morris 2007). AEA's Guiding Principles acknowledge the evaluator's responsibility to ensure non-maleficence, stating that,

> Because justified negative or critical conclusions from an evaluation must be explicitly stated, evaluations sometimes produce results that harm client or stakeholder interests. Under this circumstance, evaluators should seek to maximize the benefits and reduce any unnecessary harms that might occur, provided this will not compromise the integrity of the evaluation findings. Evaluators should carefully judge when the benefits from doing the evaluation or in performing certain evaluation procedures should be foregone because of the risks or harms. To the extent possible, these issues should be anticipated during the negotiation of the evaluation. (AEA 2004)

Reducing or mitigating unnecessary harm is extremely vital within violently divided societies. Ford et al. (2009) note that within these contexts the dissemination of sensitive findings, be they from research or evaluation, can lead to the expulsion of organisations from conflict areas or the penalisation of individuals. Evaluators must therefore balance their obligation to report the truth with their responsibility to stakeholders and participants to prevent harm that might be caused by an unfavourable evaluation.

Morris (2007) highlights that evaluators are most likely to encounter ethical conflicts during the communication of results, pressure to misrepresent evaluation results being most common.

> This pressure usually comes from the evaluation's primary client (but occasionally from the evaluator's superior), who wants the program portrayed in a more positive light (occasionally more negative) than the evaluator believes is warranted by the data. Sometimes disagreement focuses primarily on what the findings mean rather than on how positive or negative they are. (Morris 2007, 19)

The interpretation of results is an important concern for evaluations, since different evaluators could have extremely different criteria for judging what qualifies as success. For example, a psychosocial project that I evaluated in 2007 in eastern Sri Lanka could have had very different findings if the meaning of the results had not been interpreted through a critical local lens. Although the original project objectives were geared towards traditional individualistic psychosocial care, the local NGO and the community decided to use the project money to build houses for the tsunami-affected community. After field interviews, discussions with project staff members and much contemplation, I decided it was justifiable to build houses as a psychosocial project. The reason for this justification was based on the general Sri Lankan cultural ideology: 'a roof over one's head gives peace of mind', which is imbued with the idea that when there is a house, people feel better. The aim of this psychosocial project was to improve the peace of mind of tsunami-affected communities, although it had different objectives and activities in mind. Based on this perspective I could see how one could justify building houses as a psychosocial project instead of conducting activities in the original project design. Building houses was helping the community, since there was a strong community involvement and the beneficiaries of course felt better. Due to this angle of the evaluation process, there were

no issues in accessing local communities, beneficiaries and other stakeholders. It was a long and difficult negotiation with the donor, but at the end they accepted this argument. They even published the work as one of their successful tsunami projects in Sri Lanka. Further, due to the mutual understanding cultivated through this evaluation, the donor still continues to work with the local NGO and the community within a broader development agenda.

Enabling the critical local perspective to guide evaluation findings is an important consideration within contexts of violence. Evaluators who are finding it difficult to settle

> **Enabling the critical local perspective to guide evaluation findings is an important consideration within contexts of violence.**

on a single evaluation judgement may find it useful to acknowledge the multiple interpretations that the findings may bring. The African Evaluation Guidelines highlight the importance of diverse perspectives by recommending the inclusion of multiple interpretations:

> The rationale, perspectives and methodology used to interpret the findings should be carefully described so that the bases for value judgments are clear. Multiple interpretations of findings should be transparently reflected, provided that these interpretations respond to stakeholders' concerns and needs for utilization purposes. (AfrEA 2002)

While the provision of multiple interpretations may provide needed space for evaluators to articulate findings that align with the diversity of local values, evaluators should take care about adding interpretations to simply appease pressure to alter findings. Morris (2007) highlights that another common ethical challenge for evaluators is the personal and/or professional risks that doing the right thing may create. Hendricks highlighted that 'The [AEA] Guiding Principles allow me no latitude to withhold important information simply because sharing it might make my job more difficult. In fact, the Guiding Principles clearly urge me to share all relevant information without consideration of how it affects me personally. That is, however I decide to act, I should not weigh too heavily the ramification for me professionally' (Hendricks quoted in Morris 2007, 201). In violently divided societies, the requirement to divulge all information can place evaluators in danger and may have serious ramifications for their personal and professional lives.

This is why it is important that the principles of 'Do no harm' are interpreted with a critical local lens, and applied to all evaluation participants, including the evaluator. In instances where evaluators fear personal harm, they must rely on their practical knowledge, not necessarily the prescribed ethical protocols to help mitigate the situation. To apply a local lens, researchers and evaluators must take time to acknowledge the experiences of the local community and the commitments and responsibilities that their work is providing for them. They need to work with evaluation participants to interpret and comprehend ethical guidelines from a local perspective.

Knowing Your Values and Respecting the Values of Others

According to the AEA's Guiding Principles, part of cultural competence is seeking awareness of your own culturally based assumptions, and then seeking to understand the worldviews of culturally different participants and stakeholders in the evaluation (AEA 2004). Various scholars, including Humberto Reynoso-Vallejo, the Director for

Program Evaluation with the Center for Health Policy and Research at the University of Massachusetts Medical School, highlight the need for *cultural humility*, as opposed to cultural competency, because the former takes into account the political power imbalances inherent within evaluation processes.

> Cultural humility assumes that individuals' life experiences and multiple affiliations (e.g. racial/ethnic group, gender group, age cohort, region, religion, and leadership roles) interact in complex ways to shape their views. This approach assumes that the political and economic position of the group from which an individual comes, their life experiences, as well as the larger national culture shape perspectives and behavior. (Reynoso-Vallejo 2012)

Adopting a framework of cultural humility means committing oneself to ongoing self-evaluation and self-critique processes that help individuals identify their own values. House and Howe (1999) note that it is useful to think about values and facts as existing on a continuum in which brutal facts are positioned at one end and bare values at the other. Evaluative statements or claims often fall somewhere towards the centre of this continuum, where facts and values blend.

Adopting a framework of cultural humility means committing oneself to ongoing self-evaluation and self-critique, processes that help individuals identify their own values.

Scholars such as Morris also contend that it is important to acknowledge this delineation because 'personal values can influence one's response to numerous features of the project — for example, ways in which specific stakeholder groups (e.g., females, youth, the elderly, ethnic minorities, religious fundamentalists, the disabled) are treated or the degree to which one feels justified in drawing generalised conclusions from evaluation data' (Morris 2007, 200). In violently divided societies there is often a multiplicity of actors who hold different values and have played different roles: for example, those who have been the perpetrators, the victims, the bystanders, and even people who may be completely unaware of the violence surrounding them. These challenging contexts can be particularly disorienting for an evaluator who is not confident and honest about his or her moral values.

The word 'moral' can be confusing, as it can be used in two different senses. According to the *Concise Oxford Dictionary*, 'moral' is 'Concerned with goodness or badness of character or disposition' (1989, 657). Evaluators constantly negotiate and renegotiate important relations with their participants. Through engaging with the ethical principles provided by our employers, as well as what we have learnt from our religions and traditions, we create our own morals that make sense of our findings. By virtue of the evaluations we conduct, the places we travel and the people we encounter, we live according to implicitly moral bearings. However, what we frequently miss is that when we encounter people we are dealing not only with our own moral frameworks but also with theirs. We make our judgements based on our morals; they make their judgements based on their moral frameworks. In some instances, these can be harmonious, but in other situations quite conflicting. Conducting evaluations in a violently divided society can be complicated because of this unavoidable condition. However, acknowledging your own moral values is an important and necessary prerequisite that will help evaluators navigate the juxtaposition of different moral frameworks.

The journey I took to uncovering my personal values and beliefs occurred as I transitioned roles in the development field. Before I found my way into the academic and evaluation world I was a local NGO staff member in Sri Lanka in the humanitarian field. During this time I was asked to participate as a research subject in a study being conducted by universities (from both the global North and South). The researchers were interested in studying the traumas faced by humanitarian workers. Again and again, I was asked to explain the links between my experiences in a conflict-affected society and how I dealt with my traumas. There were many questionnaires, and I felt that all the questions were pushing me towards a label of 'being traumatised', while I never felt that I had a special problem different from anyone else in Sri Lanka.[7] Whenever I tried to explain this contradiction I was, and continue to be, treated as an outcast by the majority of researchers, since I do not fit into their criteria of trauma.

During my time in the field, I spent many sleepless nights thinking about what defines good and what defines bad when we apply this moral-evaluative question to our own work. We try to live our lives in ways that 'feel right' to us. We also judge people who do not appear to live the same moral lives that we think are good. In the early stages of my work, as an outsider, these juxtaposition judgements became increasingly frustrating, confusing and upsetting. As Kleinman puts it:

> That is why, in this first sense, what is moral needs to be understood as what is local, and the local needs to be understood to require ethical review [from the outside and from those on the inside who challenge accepted local values]. (Kleinman 2006, 2)

Making evaluative judgements on issues such as gender, power relations, conflict resolution and identity requires an understanding of how such issues are perceived locally, and to understand the local or insider's view requires a collaborative ethical review undertaken between the outside researcher and the inside communities. Several of the ethical frameworks emphasise that 'evaluators should be aware of different cultures, local customs, religious beliefs, gender roles... and be mindful of the potential implications of these differences when planning, carrying out and reporting on evaluations' (UNEG 2008, 14). While the failure to demonstrate cultural humility can corrupt any evaluation or research setting, it poses particular ethical challenges within violently divided societies and can lead to severe implications for both the process and the product of the study or evaluation.

Making evaluative judgements on issues such as gender, power relations, conflict resolution and identity requires an understanding of how such issues are perceived locally, and to understand the local or insider's view requires a collaborative ethical review undertaken between the outside researcher and the inside communities.

The importance of cultural humility should not be underestimated in an evaluation context. The majority of evaluation frameworks acknowledge the need for cultural competency but fail to truly engage with the politics and power dynamics inherent in the outsider-evaluator and insider-participant relationship. Evaluators need to start with where they live; but inevitably have to transcend these boundaries through processes of self-evaluation and self-critique. It is only then that the evaluator can begin to understand what is moral through a critical local lens, and thus can understand his or her limitations, as well as the unique ethical implications of the specific context. This is especially important within violently divided societies, since evaluators working in these

environments are not only challenged with understanding concepts of locality, but must also navigate the politics of vulnerability.

Critical Local Lens and Ethical Frameworks

What we can glean from the previous sections is that there are three types of disjuncture which tend to occur during the application of ethical frameworks within violently divided societies: disjunctures in *application*, *interpretation* and *between insiders and outsiders*. There are, of course, also interactions between these disjunctures. Disjunctures in application depend on how the two different professional and personal cultures understand the application. Disjunctures in interpretation are always an issue with different cultures, in translating words with different values that are embodied in culture, traditions and meaning systems. Finally, disjunctures between insiders and outsiders can complicate the process; this is another example of how a cultural differences effects the understanding of concepts. Although these are not necessarily unique to violently divided societies, there are social, political, cultural, economic and environmental factors that violence brings into these disjunctures. It is important to factor these into critical local lenses and ethical frameworks.

> There are three types of disjuncture which tend to occur during the application of ethical frameworks within violently divided societies: disjunctures *in application, interpretation* and *between insiders and outsiders.*

These disjunctures identify critical gaps within evaluation guidelines, and threaten their ability to provide guidance to evaluators, commissioners and evaluation stakeholders in violently divided societies. Evaluators who focus too narrowly on applying these professional codes without sufficient reflection are often at risk of larger ethical dilemmas because they have not recognised the other important dimensions of ethics, which acknowledge that all of these protocols must be viewed in relation to local ethical norms. This does not mean that local norms should be uncritically adopted as ethical, but rather that outsider ethical protocols and insider norms need to be reviewed together in order to determine appropriate practice for each unique evaluation context. The argument here is that in order for these frameworks to be relevant to a violently divided context it is imperative that they adopt a critical local lens. This means that a respectful, honest, transparent and accountable relationship will be built between the outside evaluator and insider community participants. These two stakeholders should be partners in ethically reviewing and acknowledging potential ethical dilemmas during the planning of an evaluation.

Applying a critical local lens is best done at the time of evaluation planning. Evaluators will benefit from early analysis of ethical considerations, since preventing ethical problems from occurring is preferable to, and often easier than, responding to problems that emerge. Morris (2007) recommends using the entry/contracting stage to think through and discuss potential ethical scenarios with stakeholders. 'The more thoroughly these matters are discussed at the beginning of the evaluation, the less likely they will arise in problematic fashion later on. And if they do arise, a framework for addressing them has at least been established' (Morris 2003 quoted in Morris 2007, 197).

Although it is not always possible to have these conversations with all stakeholders in violently divides contexts, there is still a clear benefit if some local stakeholders are

engaged in an ethical review (either formally or informally) before a study commences. This process not only helps to mitigate the emergence of ethical problems, but also helps to establish confidence in the evaluator's practical knowledge, and serves as a gentle reminder of the fundamental ethical principles that guide evaluation and research. Lastly, this process also helps to provide an opportunity for trust to be developed. Through this method, the aforementioned disjunctures can be overcome and the evaluation can be fruitful for both parties.

In many ways, this discussion has provided an opportunity to examine some of the ethical challenges encountered in violently divided societies. Apart from the personal qualities of the evaluator, there are many institutionalised problems of ethics, which can be resolved through establishing transparent and accountable ethical and methodological processes. The following three suggestions highlight several ways to improve ethical protocols so that they can be more applicable to violently divided societies:[8]

- First, we should appreciate that the existence of institutional ethical frameworks, guidelines and standards does not mean that there is agreement among actors (the evaluators and participants) about what constitutes as an ethical issue. Disagreements are common. Much work remains to be done in examining the nature of these disagreements and the strategies that might be used to address them.

- Second, it is important that evaluators prepare themselves to deal with the finale of an evaluation process. There may be problems of presentation of findings, misinterpretation and misuse of results, and/or difficulties with disclosure agreements. These problems can be avoided by establishing an honest, transparent, accountable and respectful evaluation process.

- Third, ethical frameworks, guidelines and standards should continue to be assessed systematically and periodically, and where appropriate should be reworked to reflect local values (for example, the AfrEA Guidelines). Evaluators should also understand that these ethical frameworks are part of an evolving process of self-examination by the profession within a global multicultural context and should be revisited prior to each new evaluation context to ensure they are interpreted with a critical local lens.

- Lastly, there is a strong need for meta-evaluation (that is, the evaluation of evaluations) to play an integral part in building and reshaping evaluation standards and practice. Meta-evaluation can be used by evaluators and evaluations commissions to uncover lessons from the field and hold evaluators to account.

While these suggestions are only just a start, they will help build awareness of the realities that evaluators working in violently divided societies experience. In conclusion, what is clear from the participant testimony in this article is that violently divided societies are dynamic arenas, where a variety of unique ethical dilemmas play out. If evaluation and research guidelines continue to be dominated by increasing universal bureaucratic frameworks, which follow only mainstream scientific approaches, they will be useless in these complex contexts. Applying these protocols without reflecting upon how they intersect with local realities cannot only disrupt the evaluation process, but put evaluation stakeholders, participants and even the evaluator in harm's way.

As a result, the evaluator as an 'outsider' and participant community as 'insiders' need to work together to review both institutional and community ethical frameworks, in order to

> **As a result, the evaluator as an 'outsider' and participant community as 'insiders' need to work together to review both institutional and community ethical frameworks, in order to establish a unique and effective frame of ethics for the particular evaluation project.**

establish a unique and effective frame of ethics for the particular evaluation project. This has to be an honest, transparent and accountable mechanism that maintains the integrity of both the evaluator and community. This will then become the legitimate 'ethics review committee' of the evaluative outcomes; a process that may positively transform the evaluator and community participants in the study.

> We need outside help for analysis and understanding of our situation and experience, but not for telling us what we should do.
>
> An outsider who comes with ready-made solutions and advice is worse than useless. He must first understand from us what our questions are, and help us articulate the questions better, and then help us find solutions. Outsiders also have to change.
>
> He alone is a friend who helps us to think about our problems on our own. (From a dialogue with a facilitator in the Bhoomi Sena Social Movement, Maharashtra, India 1978/79; quoted in Wignaraja 2005, 1)

Conclusion

> The question of Eurocentricism is now entirely blasé. Of course Europeans are Eurocentric and see the world from their vantage point, and why should they not? ... The question is rather the manner in which non-European thinking can reach self-consciousness and evident universality, not at the cost of whatever European philosophers may think of themselves for the world at large, but for the purpose of offering alternative (complementary or contradictory) visions of reality more rooted in the lived experiences of people in Africa, in Asia, in Latin America — counties and climes once under the spell of the thing that calls itself 'the West' but happily no more. (Dabashi 2013)

The purpose of this article is not to divide the world into the stereotype of North (bad) and South (good). As mentioned at the beginning, the article invites the reader to think about the complex nature of the challenges in applying Western ethical principles and agendas in non-Western contexts. As Dabashi (2013) argues in the quote above, a clear division exists between European and North American thinking, and non-Western thinking. This is also true for ethical principles and agendas in research and evaluation.

This article suggests that honest and equal collaboration is needed between researchers, evaluators and communities to develop effective ethical frameworks that help each other. In this process, researchers and evaluators change their ethical values while helping communities to deal with uncertainties and dangers in disaster and conflict.

Finally, this is a moral issue within the realm ethical principles. Although ethical principles are applied with an intention to do no harm, this article argues that Western ethical values do harm people in non-Western countries. Obviously, not all Western ethical principles are bad and not all non-Western worldviews are right. The challenge is

how to transform evaluation guidelines, frameworks and practices to better recognise the affected community as an equal partner of change.

DR JANAKA JAYAWICKRAMA trained in India, the US and UK, and is a social and cultural anthropologist. He is currently a senior lecturer in the Department of Geography at Northumbria University. In his current position, Janaka is the Programme Leader of the MSc in Disaster Management and Sustainable Development.

Endnotes

[1] It is important to note that the real distinction is one of paradigms of inquiry and the fact that one paradigm (generally more positivist and scientific) is the dominant tradition in academic practice in the global North and South. However, there are exceptions — appreciative enquiry, for example is the polar opposite of what I am describing in my critique of what knowledge is valued and 'ways of knowing'.

[2] Franz Fanon (Fanon, 1967) and Edward W. Said (Said, 2003) on the theoretical and multi-disciplinary nature of post-colonial theory.

[3] Elderly person from Peshawar, Pakistan, personal discussion with the author, August 1998.

[4] A farmer from conflict-affected eastern Sri Lanka, direct discussion with the author, October 2005.

[5] Community leader from El-Geneina, western Darfur, Sudan, direct discussion with the author, May 2005.

[6] A Rwandan theology professor and refugee, Lilongwe, Malawi, direct discussion with the author, August 2006.

[7] This leads to a question that is different from the question posed by the researchers: are *all* Sri Lankans traumatised?

[8] Adapted from Morris (2007).

References

AEA 2004, 'Guiding Principles for Evaluators', http://www.eval.org/publications/ GuidingPrinciplesPrintable.asp, accessed 26 February 2012.

AES 2012, 'Guidelines for the Ethical Conduct of Evaluations', www.aes.asn.au/about/code.pdf, accessed 18 March 2012.

AfrEA 2002, *The African Evaluation Guidelines*, Nairobi.

Canadian Evaluation Society 2012, 'Guidelines for Ethical Conduct', http://www. evaluationcanada.ca/site.cgi?s=5&ss=4&_lang=an, accessed 20 March 2012.

Capron, A. M. 1989, 'Human Experimentation' in Veatch, R. M. eds, *Medical Ethics*, Boston, MA: Jones & Bartlett: 125–172.

CDA 2007, 'Development Assistance and Humanitarian Aid in Conflict', Collaborative Learning Projects: Do No Harm: http://www.cdainc.com/cdawww/project_profile.php?pid=DNH &pname=Do%20No%20Harm, accessed 15 April 2011.

Concise Oxford Dictionary 1989, Oxford: Oxford University Press.

Dabashi, H. 2013, 'Can Non-Europeans Think? What Happens When Thinkers Who Operate Outside the European Philosophical "Pedigree"?', Al-Jazeera English: http://www.aljazeera. com/indepth/opinion/2013/01/2013114142638797542.html, accessed 10 April 2013.

Duggan, C. 2012, 'Show Me Your Impact: Evaluating Transitional Justice in Contested Spaces' in *Journal of Planning and Program Evaluation* 35: 1: 199–205.

Fanon, F. 1967, *Black Skin, White masks*, London: Pluto Books.

Ford, N., **Mills, E. J.**, **Zachariah, R.** & **Upshur, R.** 2009, 'Ethics of Conducting Research in Conflict Settings' in *Conflict and Health* 3: 7: 1–9.

French Evaluation Society 2003, 'Charter of Evaluation Guiding Principles for Public Policies and Programmes', http://www.sfe.asso.fr/intranet/ckfinder/userfiles/files/charter/english.pdf, accessed 20 March 2012.

House, R. & **Howe, K.** 1999, *Values in Evaluation and Social Research*, Thousand Oaks, CA: Sage.

Kleinman, A. 2006, *What Really Matters: Living a Moral Life Amidst Uncertainty and Danger*, Oxford: Oxford University Press.

Morris, M. 2007, *Evaluation Ethics for Best Practice: Cases and Commentaries*, New York: Guildford Press.

Newman, D. L. & **Brown, R. D.** 1996, *Applied Ethics for Program Evaluation*, Thousand Oaks, CA: Sage.

OECD-DAC 2010, 'Quality Standards for Development Evaluation', DAC Guidelines and Reference Series, Paris: http://www.oecd.org/dataoecd/55/0/44798177.pdf, accessed 19 December 2011.

Orb, A., **Eisenhauer, L.** & **Wynaden, D.** 2000, 'Ethics in Qualitative Research' in *Journal of Nursing Scholarship* 33: 1: 93–96.

Phillips, B. D. & **Morrow, B. H.** 2005, 'Social Science Research Needs: A Focus on Vulnerable Populations, Forecasting and Warnings', Paper presented at the Hurricane Forecast Socio-economic Workshop, Pomona, CA.

Raudonis, B. M. 1992, 'Ethical Considerations in Qualitative Research with Hospice Patients' in *Qualitative Health Research* 2: 2: 238–249.

Reynoso-Vallejo, H. 2012, 'HumbertoReynoso-Vallejo on Cultural Competence and Cultural Humility in Evaluation', AEA360: http://aea365.org/blog, accessed 2 May 2012.

Said, E. W. 2003, *Orientalism*, London: Penguin Books.

Sen, A. 1999, *Development as Freedom*, New York: Random House.

UNEG 2008, 'UNEG Ethical Guidelines for Evaluation', http://www.unevaluation.org/ethicalguidelines, accessed 25 July 2011.

Wignaraja, P. 2005, 'Fundamentals of Poverty Eradication in South Asia', Conference on Good Governance, Poverty Reduction and Community Resilience, Green Movement of Sri Lanka, and Disaster and Development Centre, UK, Negambo, Sri Lanka, 22–24 August.

Comparable and Yet Context-Sensitive? Improving Evaluation in Violently Divided Societies Through Methodology

Daniel E. Esser & Emily E. Vanderkamp

PLEASE SCROLL DOWN FOR ARTICLE

COMPARABLE AND YET CONTEXT-SENSITIVE? IMPROVING EVALUATION IN VIOLENTLY DIVIDED SOCIETIES THROUGH METHODOLOGY

DANIEL E. ESSER AND EMILY E. VANDERKAMP

Abstract

This article shows that the current stalemate in peacebuilding evaluation is due to disagreements between donor agencies, practitioners and scholar-practitioners about the necessity, appropriate level and purpose of such evaluations. It synthesises these three axes of disagreement in a theoretical framework, which is then applied to the case of evaluating reconciliation processes in violently divided societies. This application provides a clear methodological rationale for pursuing a metrics-driven, locally anchored approach to evaluating reconciliation instead of employing interpretive methods or globally standardised checklists. Realising the potential of this approach requires that donors, practitioners and researchers recast mutual expectations based on methodological rather than normative considerations.

Keywords: monitoring, evaluation, aid effectiveness, peacebuilding, reconciliation, community development

Introduction: The Stalemate in Peacebuilding Monitoring and Evaluation

The international peacebuilding community (IPC) faces an impasse with regard to the monitoring and evaluation (M&E) of peacebuilding initiatives. A decade of multilateral efforts to improve aid effectiveness (Easterly & Williamson 2011) has altered the financial and programmatic conditions under which the IPC operates; M&E has changed from being an afterthought to increasingly providing a key rationale for peacebuilding activities. As a result, IPC programme managers and analysts face pressure to 'get their act together' (Smith 2004, 1) and to demonstrate that peacebuilding programmes' underlying causal logics actually produce the intended outputs and contribute to targeted outcomes. While catalysing some methodological progress, primarily by further developing peace and conflict impact assessments (PCIAs; see Peacebuilding Centre 2011), this pressure has placed the IPC in a defensive position vis-à-vis donor agencies. Overcoming this situation in a constructive way requires that both sides rethink the necessity, level and purpose of peacebuilding M&E. This article employs methodological reasoning to provide concrete suggestions on how this can be achieved.

In 2007, the Organisation for Economic Cooperation and Development (OECD 2007, 24) lamented that while '[t]heories of peacebuilding and conflict resolution abound, each is

42

disproved as often as proved … Efforts that *appear*, on the surface, to be peacebuilding must also be held accountable for their actual effects in a particular context at a particular time' (ibid., 24–26, emphasis in original). In a subsequent report (OECD 2008, 10), the organisation urged donor agencies to 'promote the systematic use of evaluation for all conflict prevention and peacebuilding work, and require implementing partners, such as NGOs, to conduct evaluations'. Along the same lines, the United Kingdom's Department for International Development (DFID) emphasised 'the case for a well-structured approach to M&E in situations of fragility and conflict', which, according to the world's second-largest donor agency, 'is just as pressing, if not more so, as in other development contexts' (DFID 2010, 1–2). Similarly, the most financially invested bilateral donor supporting peacebuilding efforts worldwide, the US federal government, explicitly states that this focus on evidence-based peace programming will 'inform U.S. foreign policy and development goals' (Department of State 2012) and is 'an essential component to effectively implementing diplomatic and development programs and initiatives' (United States Agency for International Development [USAID] 2011a). Such high-level commitments to evaluating peacebuilding programmes should be operationalised by 'establishing new requirements for performance evaluations, designing rigorous impact evaluations, linking evaluations to future funding decisions, and promoting the unbiased appraisal of programs and the full disclosure of findings' (Department of State & USAID 2010).

Scholars working at the intersection of academia and practice have arrived at similar conclusions. First and foremost, they have argued that implementers must become more deliberate and rigorous in the ways in which they approach M&E (Gibson 2006; 2007; Meierhenrich 2008; Neufeldt 2011; Scharbatke-Church 2011). Indeed, according to a recent report written for the United States Institute for Peace (USIP), 'good progress has been made on the intellectual front. There are now clear guidelines, frameworks and tool kits to guide practitioners who wish to initiate an evaluation process within the peacebuilding field' (Blum 2011, 1). To face the reality of mounting expectation to demonstrate impact, as well as continued competition for external funding among IPC organisations, an embrace of comprehensive M&E frameworks would appear to be a logical response (within what is possible, given the commonly higher cost of M&E in violently divided societies). Arguably, this would not only serve organisations' interest in safeguarding their survival and growth (DeCarlo & Ali 2010) but would also reflect their often-stated commitment to learning (Neufeldt 2011). And, yet, IPC organisations' reactions to these recent demands for more comprehensive, systematic and meaningful M&E have, in fact, been mixed. The aforementioned USIP report, for instance, laments that conceptual advances have informed the practice only to a very limited extent and that 'progress in improving peacebuilding evaluation itself has slowed over the past several years' (Blum 2011, 1). Further, Scharbatke-Church (2011, 471, 475) concludes a review by arguing that 'accountability has remained nearly a non-issue in the peacebuilding field' and urges practitioners to 'make [M&E] an opportunity for accountability and learning'. The OECD (2008, 13) even went as far as accusing IPC practitioners of 'sometimes resist[ing] evaluation', insisting that 'this resistance to [M&E] must be overcome' (ibid.).

To date, several arguments have been fielded to explain the limited application of conceptual innovations in IPC M&E practices. Most prominently, practitioners and scholar-practitioners in this field have posited that the complexity of the task at hand — nothing less than building peace — cannot possibly be captured in neat causal chains following the tradition of logical framework approaches (LFA) and its agency-specific adaptations (Leonhardt 2003; Mika 2002). The hope of attributing a peaceful outcome to

43

any particular programme is futile, according to authors taking this position, since peace processes frequently transcend spatial levels, involve numerous actors and build on multiple initiatives occurring simultaneously. Moreover, peacebuilding is rendered a particularly challenging case for M&E because it aims to trigger subtle attitudinal changes within individuals and communities, such as increased trust or tolerance (Anderson et al. 2004; Lederach et al. 2007). In addition, it has been argued that focusing on the evaluation of peacebuilding rather than programme implementation risks diverting scarce financial and human resources from the core task (Bell 2000; Gasper 2000; Kelman 2008). Furthermore, since peace and reconciliation processes in violently divided societies tend to be slow-moving, short donor timeframes are inept at capturing their results (Lederach et al. 2007), thus risking that M&E be reduced to a mere 'tick-box exercise' (Hamber & Kelly 2008, 16) as part of an ongoing 'technocratisation of peacebuilding … magnified … by the nature of contemporary peace-support interventions' (Mac Ginty 2013, 57).

We aim to diffuse some of these tensions in this article by demonstrating that methodological rather than normative arguments hold the key to meaningful and effective peacebuilding M&E. In order to do so, we capture the state of the debate in three axes that help us structure the different arguments made. The first axis focuses on different interpretations of the value of M&E in the context of peacebuilding. This axis juxtaposes positions informed by the IPC's normative underpinnings, which relativise both the need and the possibility of measuring peacebuilding outputs and outcomes, with those which demand 'objective' approaches to M&E (see also Grofman 1997, 75–79). As these positions reflect constructivist and positivist (or realist, as in the case of Van Evera [2003]) stances, respectively, we label this axis 'epistemological'. The second axis captures the 'spatial' dimension, asking the question: at what level should we measure reconciliation? This axis contrasts the arguments in favour of locally specific measures of reconciliation with those urging for general indicators in order to allow for cross-national comparisons. The local/global dimension of peacebuilding M&E has been explored elsewhere (e.g., Mac Ginty 2013), though in isolation from the two other dimensions examined here. The third axis zeroes in on the 'intentional' dimension, which is divided between proponents of M&E primarily as a basis for external accountability and those arguing that it should first and foremost create opportunities for organisational learning and programme development. In other words, the intentional axis poses the question: for what purpose should we measure?

> **Methodological rather than normative arguments hold the key to meaningful and effective peacebuilding M&E.**

Following our analysis of the contemporary debate, we select the M&E of reconciliation initiatives in violently divided societies as a 'critical case' (George & Bennett 2004) in the field of peacebuilding. This field can be characterised as comprising four general aspects: socio-economic foundations, security, political frameworks, and reconciliation and justice (OECD 2007). Agreeing with Stover & Weinstein (2004), we argue that, since reconciliation represents the most challenging aspect in terms of indicator design and measurement, our recommendations are likely to be relevant to the other three aspects of peacebuilding as well. Further, reconciliation has been flagged as the least operationalised of the four aspects of peacebuilding, thus necessitating further study (Gibson 2006; 2007). We therefore conduct an appraisal of the salience of each axis of debate in the case of monitoring and evaluating reconciliation initiatives and programmes. In this section we also demonstrate why a revised approach to the M&E of reconciliation, as well as to peacebuilding more broadly, is both necessary and feasible.

44

Three Axes of Debate in Peacebuilding M&E

The epistemological axis: is objective measurement possible?

Facing growing pressure from national stakeholders to demonstrate value for money, multilateral as well as bilateral donor agencies have publicly stated their commitment to 'corrective action plans to develop customised impact indicators and strengthen monitoring and evaluation of conflict mitigation activities', as a recent audit of USAID (2010; cf. OECD 2012) points out. Among other effects, such as increasing reporting requirements, this stance has resulted in increasing technical sophistication in the methods employed to measure the impact of externally financed IPC activities, in particular a rise in mixed-method evaluations 'using a quasi-experimental design to contrast treatment and comparison communities' (USAID 2011b, 8). In order to facilitate such methods, agencies have urged that 'a number of conditions should be in place before an evaluation process begins. The most essential elements [include] clear and measurable objectives; a testable programme logic and theory of change; and monitoring tools, including performance information and indicators (in order to measure achievements on the way)' (OECD 2008, 24). Further, '[s]electing the correct indicators from the start is essential to identifying both positive and potentially negative impacts of interventions' (DFID 2010, 8). Evaluation of externally supported programmes and projects is thus characterised as a central component of international assistance in conflict-ridden societies.

IPC scholar-practitioners, on the other hand, have responded to these demands with scepticism. Wachira (2001 cited in Leonhardt 2003), for instance, reported on a case in which donors sought objective measurements while the members of a local peacebuilding initiative struggled to translate their notion of peacebuilding as qualitative, 'liberating and humanizing change' into an assessable framework. Experiences like these illustrate a widely held conviction among members of the IPC that peacebuilding has inherent value irrespective of a singular project's outputs or outcomes: 'A sustained peace process has value of its own', according to Galama and van Tongeren (2002, 21), and '[t]he focus [of M&E] thus should be on the actual process towards peace rather than just on the result, however, peaceful' (ibid.). Advocating 'collaborative and elicitive' approaches to evaluating such processes in order to 'serve as a catalyst for transforming relationships of power', Mika (2002, 339) seconded this position by drawing a line between 'conventional and staid evaluation practices that are technical in nature and actuarial in intent' (ibid.) and those presumably more suited to the M&E of IPC activities. Similarly, Anderson and colleagues (2004, 14) pointed to the difficulty of 'know[ing] whether or when a particular program outcome is significant for peace [given that] the goal of just and sustainable peace is so grand, and progress toward it immeasurable in its multitude of small steps'. More recently, Neufeldt (2011, 501) has lamented that, while the shift towards monitoring and evaluation in peacebuilding programmes is 'likely due in good part to increased integration of conflict transformation and peacebuilding into development work', 'the business model of development, with its emphasis on efficiency and results-based management [can] undermine relationships … and convey messages of disrespect to local communities' (ibid., 489).

These two diverging stances related to the value of M&E in peacebuilding illustrate a fundamentally different perception of the sources of legitimacy for externally supported peacebuilding programmes. Whereas donor agencies widely posit evaluation as a legitimising activity both vis-à-vis beneficiaries and in the face of mounting public scrutiny, IPC practitioners have expressed concern that a greater emphasis on evaluation rather than implementation — despite M&E's commonly holistic conception — could

45

potentially undermine the very legitimacy of their activities within beneficiary communities. These practitioners do not necessarily reject evaluation as such and may in fact agree that evaluations contribute to the overall mission, but they are opposed to approaches that, they fear, are not suited to adequately capturing place-dependent interpersonal dynamics and may therefore alienate local partners via elitist discourses and sophisticated — commonly quantitative — methods. Indeed, the hyperlocality of peacebuilding efforts is so prominent a theme in both IPC literature and practice that it requires separate analysis. Here, too, international agencies have been demanding a change in approach.

Two diverging stances related to the value of M&E in peacebuilding illustrate a fundamentally different perception of the sources of legitimacy for externally supported peacebuilding programmes.

The spatial axis: context-specific versus globally comparable metrics[1]

Conflict transformation requires context-sensitive programming. One of the most prolific authors in this field has, for over a decade, reminded practitioners and aid planners that, '[t]o be at all germane to contemporary conflict, peacebuilding must be rooted in and responsive to the experiential and subjective realities shaping people's perspectives and needs' (Lederach 1997, 24). Galama and van Tongeren (2002, 18) concur, arguing that '[t]he process of peacebuilding should begin with the people who are affected by conflict, with their experiences, questions and their own experiences towards peacebuilding'. In response, Scharbatke-Church (2011) has urged programme implementers to seek community input into monitoring and evaluation not only during the final stages of evaluation but also throughout the design and implementation process.

Although there is no doubt about the need for locally meaningful approaches to peacebuilding, the critical question that arises in the context of contemporary peacebuilding evaluation — for instance, during a roundtable session at the 2013 Convention of the International Studies Association (ISA 2013) — is whether programmes that respond primarily to local needs can, in fact, be evaluated by applying a set of metrics that retains validity beyond the local level. Diehl and Druckman (2010, 8) make this point forcefully by questioning the usefulness of context-dependent metrics:

> Case-specific standards or indicators inhibit the ability of policymakers to take what they learned from one operation and adapt that to a different context. From a scholarly standpoint, researchers must be able to construct some common standards and indicators of success in order to compare performance across missions and to draw generalizations. Case-specific benchmarks inhibit the empirical verification of propositions and theories about peace operations and thereby stifle the development of general knowledge and patterns.

Leading bilateral and multilateral agencies such as USAID, the State Department and the United Nations have embraced this position by issuing detailed handbooks listing hundreds of indicators to be used at the programme and national levels in order to enable cross-national performance measurement (USAID 2011c; Department of State 2011; UN 2010; 2011). Yet reconciling locally meaningful measures and metrics with generalisability appears problematic. Although the quest for globally valid metrics is understandable from a political perspective, to prove meaningful in practice, peacebuilding evaluation must rely on locally validated measures. However, it is not clear how a compromise

between these two positions would not jeopardise one of the core strengths of the IPC community — a deep understanding of local conflict dynamics — while also failing to enable comparability across settings with different conflict trajectories.

The intentional axis: why engage in evaluation?

The third axis of debate between the IPC community and international donor agencies refers to different priorities with regard to the primary purpose of evaluating peacebuilding activities. Although less of a dichotomous setup than in the cases of the previous two axes of debate, donors on the one hand and the IPC on the other commonly attach varying degrees of importance to intra-organizational and IPC-wide learning in the context of peacebuilding M&E, which in turn is a function of their distinct accountability relations. Lederach et al. (2007, 57) have warned that accountability and learning may in fact constitute a trade-off. It is not that donors dismiss learning per se, nor do implementers generally undervalue accountability. However, for donors, reporting tends to take precedence over reflection, since reports rather than reflection ensure continuous budget allocation. As a result, among donors '[i]nstitutional learning, which is reflecting on the lessons learned from particular programmes, is undertaken on an irregular basis' (Leonhardt 2003, 56). In Galama and van Tongeren's (2002, 23) words, '[d]ifferent actors have different needs when it comes to evaluation'. Whereas donor agencies conduct and utilise evaluations primarily to forge accountability from implementers and to respond to their own bureaucratic accountability relations, practitioners see evaluative data and findings primarily as part of

> **Donors on the one hand and the IPC on the other commonly attach varying degrees of importance to intra-organizational and IPC-wide learning in the context of peacebuilding M&E, which in turn is a function of their distinct accountability relations.**

reflective practice. Indeed, despite the most recent OECD report (2012) on peacebuilding evaluation positing that a synthesis between accountability and learning is possible, systemic incentives remain stacked against it: each of the two relative emphases is based on a different interpretation of the same perceived problem, i.e., the underperformance of peacebuilding as a key area of international cooperation. In response to this underperformance, and mindful of mounting political pressure, donors prioritise the maximisation of both external (i.e., from IPC implementers) and internal (i.e., within their own polities) accountability. Conversely, organisations in charge of running peacebuilding programmes on the ground are keen on improving operations while strengthening ties with beneficiaries, yet without collecting and publicising data that, if pointing out problems or frictions, may actually endanger their financial survival.

Unsurprising in light of these systemic dynamics, leading IPC scholar-practitioners have advocated a 'special emphasis on monitoring- and evaluation-as-learning, *rather than* evaluation-as-measuring results' (Lederach et al. 2007, 2, emphasis added). In order to operationalise this objective, they envision '[l]earning communities [involving] various circles of people, depending on the purpose of the learning event' (ibid., 8). This approach stands in stark contrast with the OECD's (2008, 10) depiction of evaluation as supporting learning and accountability simultaneously, and it is precisely this emphasis on generalisability that presents a multifold challenge to the IPC community, as we have outlined in the previous two subsections.

We visualise the main findings from our review in Figure 1. Although the depiction is necessarily stylistic and cannot capture the myriad of positions and approaches prevalent

Figure 1: Three dimensions of the debate on peacebuilding M&E and positions of key stakeholders.
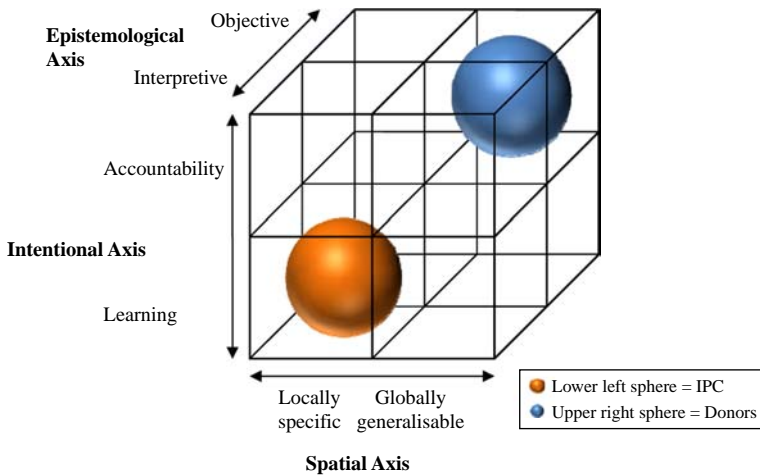
in peacebuilding M&E today, the figure helps us better understand the sources of tension as well as the potential for rapprochement between key stakeholder groups.

Figure 1 visualises the IPC's continued reliance on local approaches utilising interpretive measures in an attempt to maximise contextual validity while international donors increasingly seek globally generalisable quantitative indicators. In the following section, we investigate how each of these three axes of debate has come to bear on the monitoring and evaluation of reconciliation initiatives and programmes in violently divided societies. We then show that revisiting and rethinking some of the positions that currently frame the evaluation of reconciliation programmes offer an opportunity to overcome the stalemate and measure and subsequently improve aid effectiveness in peacebuilding.

Evaluating Reconciliation in Violently Divided Societies

The case of M&E of reconciliation processes allows us to test the salience of each of the three aforementioned axes — epistemological, spatial and intentional — in what Gibson (2006, 2007) has flagged as the most under-researched component of peacebuilding. In his research on the South African Truth and Reconciliation Commission (TRC), Gibson (2006, 90) asserts that 'concepts like "truth" and "reconciliation" can be (and should be) measured and assessed using rigorous and systematic social science methods', thus seemingly embracing the positivist stance held by international agencies. In order to measure movement towards sustainable peace, Gibson (2007, 257, emphasis added) defines reconciliation as 'the development of some *means* by which those who were previously foes (and even co-combatants) can agree to coexist and compete peacefully rather than violently'. Gibson proposes the use of a Reconciliation Barometer to follow changes in reconciliation over time (ibid., 278). Similarly, Kelman (2004, 124) suggests that acknowledgement of responsibility should not be merely symbolic. Rather, he argues that acknowledgement must be validated through 'compensation, reparation, and restitution'. Although Gibson (2007, 260) and Kelman (2004, 124) explicitly acknowledge the importance of beliefs and attitudes, their emphasis lies on measuring

48

the concrete actions taken to facilitate coexistence and peaceful political as well as economic competition.

Arguing from a constructivist perspective, Brounéus (2008) disagrees with this conceptualisation and its resultant operationalisation. Her definition of reconciliation focuses on a 'societal process that involves mutual acknowledgement of past suffering and the changing of destructive attitudes and behaviors into constructive relationships toward sustainable peace' (ibid., 294). Seen from this perspective, perceptional rather than action-oriented indicators must remain central in M&E. Dziedzic and colleagues (2008, 52–53) follow this line of argument by suggesting several indicators aimed at assessing changes in perceptions among adversarial groups, such as public satisfaction in the way past abuses have been dealt with, the extent to which perpetrators have acknowledged past wrongs, the extent to which victims have forgiven perpetrators, and the degree of tolerance and readiness to compromise with members of other identity groups.

Lederach (1997) offers yet another way of envisioning reconciliation which underscores the normative tenets of the IPC. He conceives reconciliation as both a 'focus' — reconciliatory processes build relationships between antagonists — and a 'locus' — successful reconciliation 'creates a space for encounter by the parties, a place where the diverse but connected energies and concerns driving the conflict can meet, including the paradoxes of truth and mercy, justice and peace' (ibid., 34–35). Finally, Villa-Vicencio (2009, 151) advocates for a clear distinction between 'individual' reconciliation and 'national' reconciliation. Although this proposition resonates with our second axis of debate (discussed above and applied below), Villa-Vicencio's main motivation is to distinguish psychological from political aspects of reconciliation. Overall, he characterises reconciliation as 'both *process* and *goal*' (ibid., 170, emphasis in original), the former being 'inevitably uneven' while the latter implies that 'people have equal access to essential social services and basic material necessities' (ibid.). Ultimately, he argues, 'reconciliation is an art rather than a science' (ibid., 171).

This brief comparison of approaches demonstrates that there is wide disagreement regarding the definition of reconciliation. Yet, how one defines reconciliation influences how programmes are monitored and evaluated. Aside from illustrating an impressive diversity of conceptualisations, these different definitions suggest that one of the reasons why appropriate approaches to the M&E of reconciliation in violently divided societies have remained vague is the 'lack of clarity about what, specifically, reconciliation mean[s]' (Hamber & Kelly 2008, 7). Notably, Meierhenrich (2008, 196) even charged the IPC with conceptualising and operationalising reconciliation 'with insufficient rigor' as a result of 'conceptual stretching' (ibid., 204) of

Different definitions reflect varying epistemological positions on reconciliation, often connected to normative agendas, which in turn shape different approaches to M&E.

the term in order to fit a wide variety of settings and to thus increase the relevance of peacebuilding as a field. More fundamentally, these examples demonstrate that different definitions reflect varying epistemological positions on reconciliation, often connected to normative agendas, which in turn shape different approaches to M&E.

Spatially, research on the effectiveness of reconciliation processes in violently divided societies has commonly criticised donor-supported national reconciliation programmes as insufficiently contextualised and, in response, has measured effects locally rather than at the national level. For instance, Hermann (2004, 44) has questioned the utility of

programmes focussed at the state level, suggesting instead that 'the relevant unit of analysis may in fact be smaller than the "society" at large'. Similarly, Shaw (2005, 4–7) has challenged the idea of 'national healing' in her work on the Sierra Leonean TRC, positing instead that 'local practices of healing and reintegration' (ibid.) mattered most. Shaw argues that, while a TRC may have been an appropriate vehicle to establish collective memory in South Africa, cultural and contextual differences diminished the effectiveness of the national TRC in Sierra Leone. Her proposition is backed by research carried out by Smith (2010, 55, 57), who detected public wariness of the state-led Sierra Leonean TRC, often resulting in participation in community-driven models informed by a 'forgive and forget mentality' instead. Analogous evidence has also been reported from other African settings. Contrasting the case of Rwanda's nationally orchestrated reconciliation process with Mozambique's grassroots approach, Brounéus (2008, 306–307) found support for her hypothesis that community-level reconciliation efforts that stress attitudinal change are more effective than national campaigns.

Echoing Scharbatke-Church's (2011) related concerns, reflections on the relationship between learning and accountability in monitoring and evaluating reconciliation initiatives have been few and far between. Learning is commonly defined narrowly, namely, as the participatory development of indicators; those authors who mention accountability in the context of reconciliation do so under a different pretext. For instance, by advocating for 'an appropriate balance between accountability and human rights on the one hand and peace and reconciliation on the other', Villa-Vicencio (2009, 32–33) adheres to a conceptualisation of accountability that is judicial rather than organisational.

However, this does not imply that learning and accountability do not matter to sponsors and implementers of reconciliation initiatives and programmes, albeit to varying degrees. In the only publication discussing inter-organisational accountability in the context of reconciliation processes in violently divided societies, Neufeldt (2011) argues that it should be directed first and foremost to those the programmes are meant to serve. Here again, the normative underpinnings of the IPC come to the fore and risk obfuscating the procedural requirements for complex accountability mechanisms to work. While practitioners commonly favour internal learning and local accountability, donors want to objectively assess the impacts of reconciliation programmes, both positive and negative. This paradigmatic rift between practitioners and donors gets to the core challenge of M&E of reconciliation initiatives. Both in the specific case of reconciliation as well as in peacebuilding more broadly, the normative argument emphasising the need for accountability to local beneficiaries can be employed to relativise the lack of M&E rigour, thus shirking accountability to donors. In the section that follows, we develop a methodological rationale for metrics-driven, locally validated approaches, which we argue are best suited to ensuring external accountability while also allowing for procedural and programmatic learning.

Toward Better Evaluation: The Logics of Convergence

We begin our process of rethinking M&E in violently divided societies with the seemingly least contentious axis of debate. Although donor agencies' common interest in nationally aggregated and cross-nationally comparable data is both politically and methodologically comprehensible, our review of prevailing arguments generates strong support in favour of locally formulated indicators. In fact, Diehl and Druckman's (2010) plea for global indicators is methodologically callow. Its underlying assumption that global indicators are

most suited to ensuring comparability is a fallacy because the meanings of most of these indicators are likely to differ across countries and even within countries. At the same time, such an ex ante rejection of general indicators does not, as they suggest, render comparison impossible per se. Whenever local consultations in different settings happen to produce substantively matching indicators, comparing the data on these indicators is entirely possible and indeed analytically desirable. Sequence thus becomes a central concern; local validation of indicators should precede cross-national comparisons in order to ensure that the latter are based on *conceptually* comparable measurements. Methodologically, this implies that meta-evaluations across programmes that seek to establish general knowledge must be based on consultations with those under study in order to ascertain that data on locally defined indicators allow for valid comparison with data from other sites even if indicators themselves are worded identically: meaning, not wording, matters.

> **Meta-evaluations across programmes that seek to establish general knowledge must be based on consultations with those under study in order to ascertain that data on locally defined indicators allow for valid comparison with data from other sites even if indicators themselves are worded identically: meaning, not wording, matters.**

As Hamber and Kelly (2008, 7) rightly point out, confusion about the meaning of reconciliation among respondents invariably invalidates their responses analytically. Following this logic, assessments of reconciliation must be driven by local participation. A shared understanding of reconciliation at the community level is a crucial first step for producing valid M&E data. As locally developed indicators do not preclude comparisons of data from different settings, they serve to ensure internal validity by prompting local communities to define appropriate terminology. Any subsequent comparison of local data must be mindful of such terminological contexts in order to preserve external validity. At the same time, it becomes obvious that no handbook or tool kit offered by donor agencies in the interest of reaping economies of scale through purportedly clean data can ever produce meaningful M&E results if it does not contain an annex detailing both local validation strategies and empirical results.

If local ex ante consultation results in a commonly shared definition of reconciliation, then rigorous data collection becomes a joint responsibility. Donor agencies thus have a point as well: as much as the IPC has been cherishing its anti-realist foundations, methodological weaknesses in both theory and practice cannot be justified on normative grounds. There is no basis on which to argue that context-appropriate surveys — e.g., using social distance and Likert scales as well as multiple choice and comment options — cannot be employed meaningfully to systematically assess changes in attitudes and beliefs at the micro level (Denskus 2012; McIntosh 2008; Tiemessen 2008). Such methods can, for instance, elucidate the extent to which both victims and perpetrators feel the initiative was just and even-handed. Moreover, using these methods to gauge perceptions of changing interpersonal relations and procedural fairness (Leach & Sabatier 2005) can, either concurrently or at a later stage, be integrated into strictly quantitative analyses of changes in the incidence of violence and their causal logics. The same is true for quasi-experimental designs, as long as their design is conceptually validated locally and as part of programme design. In short, relativism simply has no defensible grounds in peacebuilding M&E.

This methodological rationale for rethinking M&E in peacebuilding prompts both donor agencies and the IPC to reconsider current demands and expectations. Donors have to come to terms with methodological limitations to scaling up M&E, but they can expect

Figure 2: Spaces of convergence in peacebuilding M&E.

Locally anchored approaches to evaluating reconciliation gain further momentum by addressing the seemingly intractable trade-off between internal learning and account-ability in peacebuilding practice. Systematically collected data towards locally defined measurement frameworks as outputs of participatory processes maximise opportunities for learning among all stakeholders, including those who supposedly benefit from the activities undertaken. At the same time, they strengthen the position of implementing agencies vis-à-vis international agencies, since they enable the former to generate and report valid data. In turn, receiving such data lies in the interest of agencies as they allow for a more valid results-oriented assessment of international assistance. Figure 2 summarises these arguments.

Conclusion: The Case for Cooperation

Our proposition of a convergence between donors and the IPC on a locally validated, rigorous approach to peacebuilding M&E is not based on wishful thinking. Throughout this article, we have acknowledged and indeed embraced the political economy of aid and its resulting incentive structures. We hope that, on this basis, we have made a case for why methodological — rather than normative — arguments, if taken seriously, can provide useful guidance for rethinking current positions in peacebuilding M&E and for acting collaboratively to achieve shared objectives. In addition, we have shown that the persistent lack of definitional clarity concerning reconciliation — a commonly cited cause of

implementing partners to improve their capacity to produce or support community-level research that meets scientific quality standards. IPC practitioners, in return, need to ask themselves whether it is ethically defensible to let the field's normative underpinnings limit its possibilities to improve the quality of its work. At the same time, IPC practitioners have a compelling case for retaining the field's tried-and-tested focus on engaging people in the communities in which they live and to do so in participatory ways that build meaningful relationships.

> **Donors have to come to terms with methodological limitations to scaling up M&E, but they can expect implementing partners to improve their capacity to produce or support community-level research that meets scientific quality standards.**

imprecise measurement — fails to justify inaction, since definitional clarity can, in fact, be achieved situationally through participatory methods. Given that data generated by such approaches are unavoidably site-specific, they render global frameworks meaningless unless these are compiled anew for each comparative exercise, which is rarely the case. This, then, constitutes the true trade-off that donor agencies need to accept: the quest for 'objective' measurement can be satisfied only within a local context, and comparison is meaningful only if indicators thus validated align substantively. Agencies that are politically willing and able to accept this logic and its implications for practice can and should expect their implementing partners to conduct local evaluations that are in line with scientific quality standards. At the same time, implementers are presented with an opportunity to move towards more constructive engagement and dialogue with donors. Budgeting must be a central component of this renewed conversation; implementing organisations are right to expect donors to back up demands for rigorous evaluations financially. Finally, locally deliberated and validated indicators also maximise the potential for learning among and accountability among IPC stakeholders at different levels. We know that it will take time for a rapprochement between agencies and implementers to happen. Yet we believe that if agencies and the IPC are serious about their commitments to supporting non-violence globally, both camps will eventually realise that there are simply no workable alternatives to scientific logic.

DANIEL E. ESSER is Assistant Professor of International Development at American University in Washington, DC. He has carried out research in Afghanistan, Pakistan, Sierra Leone and Mexico and has published on the political economy of post-war reconstruction, agenda-setting in international assistance and global development discourses. He is double graduate of the London School of Economics and Political Science and worked full-time for the International Labour Organisation (ILO), the United Nations Development Programme (UNDP) and the UN Economic and Social Commission for Asia and the Pacific before rejoining academia.

EMILY E. VANDERKAMP graduated from American University in 2012 with a master's in ethics, peace, and global affairs. Her main interests include conflict resolution, programme management, and the role of women in development, with a regional focus on the Middle East and North Africa. She also holds a BA degree (summa cum laude) from the College of William and Mary.

Endnote

References

Anderson, M.B., Chigas, D., Olson, L. & Woodrow, P. 2004, *Reflecting on Peace Practice Handbook*, Cambridge, MA: Collaborative for Development Action. www.cdainc.com/rpp/docs/ ReflectingOnPeacePracticeHandbook.pdf

Bell, S. 2000, 'Logical Frameworks, Aristotle and Soft Systems: A Note on the Origins, Values and Uses of Logical Frameworks, in Reply to Gasper' in *Public Administration and Development* 20: 1: 29–31.

Blum, A. 2011, *Improving Peacebuilding Evaluation: A Whole-of-Field Approach*, Washington: USIP.

Brounéus, K. 2008, 'Analyzing Reconciliation: A Structured Method for Measuring National Reconciliation Initiatives' in *Journal of Peace Psychology* 14: 3: 291–313.

DeCarlo, A. & Ali, S.H. 2010, 'Developing a Reconciliation Indicator', 19 January: http://www.policyinnovations.org/ideas/innovations/data/000154

Denskus, T. 2012, 'Challenging the International Peacebuilding Evaluation Discourse with Qualitative Methodologies' in *Evaluation and Program Planning* 35: 148–153.

Department of State. 2011, *Conflict Mitigation and Reconciliation* [indicators], Washington: United States Department of State. http://www.state.gov/documents/organization/101762.pdf

Department of State. 2012, *Program Evaluation*, Washington: United States Department of State. http://www.state.gov/s/d/rm/rls/perfrpt/2010/html/153549.htm

Department of State & USAID 2010, 'Leading through Civilian Power: The First Quadrennial Diplomacy and Development Report', Washington: http://www.state.gov/documents/organization/153635.pdf

DFID 2010, 'Briefing Paper 1: Monitoring and Evaluation, Working Effectively in Conflict-Affected and Fragile Situations', London: http://www.dfid.gov.uk/Documents/publications1/governance/building-peaceful-states-I.pdf

Diehl, P.F. & Druckman, D. 2010, *Evaluating Peace Operations*, Boulder, CO: Lynne Rienner.

Dziedzic, M., Sotirin, B. & Agoglia, J. 2008 eds, 'Measuring Progress in Conflict Environment (MPICE): A Metrics Framework for Assessing Conflict Transformation and Stabilization', Working Paper, United States Institute of Peace.

Easterly, W. & Williamson, C.R. 2011, 'Rhetoric Versus Reality: The Best and Worst of Aid Agency Practices' in *World Development* 39: 11: 1930–1949.

Galama, A. & van Tongeren, P. 2002 eds, 'Towards Better Peace Building Practice' in Galama, A. & van Tongeren, P. eds, *Towards Better Peacebuilding Practice: On Lessons Learned, Evaluation Practices and Aid & Conflict*, Utrecht: European Centre for Conflict Prevention: 13–41.

Gasper, D. 2000, 'Evaluating the "Logical Framework Approach" towards Learning-Oriented Development Evaluation' in *Public Administration and Development* 20: 1: 17–28.

George, A.L. & Bennett, A. 2004, *Case Studies and Theory Development in the Social Sciences, BCSIA Studies in International Security*, Cambridge, MA: MIT Press.

Gibson, J.L. 2006, 'Overcoming Apartheid: Can Truth Reconcile a Divided Nation?' in *Annals of the American Academy of Political and Social Science January* 603: 1: 82–110.

Gibson, J.L. 2007, '"Truth" and "Reconciliation" as Social Indicators' in *Social Indicators Research* 81: 2: 257–281.

Grofman, B. 1997, 'Seven Durable Axes of Cleavage in Political Science' in Monroe, K.R. eds, *Contemporary Empirical Political Theory*, Berkeley, CA: University of California Press: 75–86.

Hamber, B. & Kelly, G. 2008, 'The Challenge of Reconciliation: Translating Theory into Practice' in *Northern Ireland Community Relations Council, A Sustainable Peace? Research as a Contribution to Peace Building in Northern Ireland*, Belfast: 3–26.

Hermann, T. 2004, 'Reconciliation: Reflections on the Theoretical and Practical Utility of the Term' in Bar-Siman-Tov, Y. eds, *From Conflict Resolution to Reconciliation*, Oxford: Oxford University Press: 39–60.

ISA 2013, 'Engaging with the New Deal: Amplifying Voices from Fragile and Conflict Affected Countries', Peace Studies Roundtable, International Studies Association Annual Convention, San Francisco, 3 April.

Kelman, H.C. 2004, 'Reconciliation as Identity Change: A Social-Psychological Perspective' in Bar-Siman-Tov, Y. eds, *From Conflict Resolution to Reconciliation*, Oxford: Oxford University Press: 111–124.

Kelman, H.C. 2008, 'Evaluating the Contributions of Interactive Problem Solving to the Resolution of Ethnonational Conflicts' in *Peace and Conflict* 14: 1: 29–60.

Leach, W.D. & **Sabatier, P.A.** 2005, 'To Trust an Adversary: Integrating Rational and Psychological Models of Collaborative Policymaking' in *American Political Science Review* 99: 4: 491–503.

Lederach, J.P. 1997, *Building Peace: Sustainable Reconciliation in Divided Societies*, Washington: United States Institute of Peace.

Lederach, J.P., **Neufeldt, R.** & **Culbertson, H.** 2007, *Reflective Peacebuilding: A Planning, Monitoring, and Learning Toolkit*, Notre Dame, IN: Joan B. Kroc Institute for International Peace Studies.

Leonhardt, M. 2003, 'Towards a Unified Methodology: Reframing PCIA' in Austin, A., Fischer, M. & Wils, O. eds, *Peace and Conflict Impact Assessment: Critical Views on Theory and Practice*: 53–66 http://www.berghof-handbook.net/documents/publications/dialogue1_leonhardt.pdf

Mac Ginty, R. 2013, 'Indicators +: A Proposal for Everyday Peace Indicators' in *Evaluation and Program Planning* 36: 1: 56–63.

McIntosh, I.S. 2008, 'Measuring Reconciliation When There Is No Roadmap: A Study of Trajectory', paper presented at International Peace Research Association Conference, Institute for International and European Policy, Leuven, Belgium, 15–19 July.

Meierhenrich, J. 2008, 'Varieties of Reconciliation' in *Law & Social Inquiry* 33: 1: 195–231.

Mika, H. 2002, 'Evaluation as Peacebuilding?: Transformative Values, Processes, and Outcomes' in *Contemporary Justice Review* 5: 4: 339–349.

Neufeldt, R.C. 2011, '"Frameworkers" and "circlers" — Exploring Assumptions in Impact Assessment' in Austin, B., Fischer, M. & Giessmann, H.J. eds, *Advancing Conflict Transformation: The Berghof Handbook II*, Opladen, Germany: Barbara Budrich: 483–504.

OECD 2007, 'Encouraging Effective Evaluation of Conflict Prevention and Peacebuilding Activities' in *OECD Journal on Development* 8: 3: 7–106.

OECD 2008, 'Guidance on Evaluating Conflict Prevention and Peacebuilding Activities: Working Draft for Application Period', Paris.

OECD 2012, *Evaluating Peacebuilding Activities in Settings of Conflict and Fragility: Improving Learning for Results*, Paris.

Peacebuilding Centre 2011, *Peace and Conflict Impact Assessment (PCIA) Handbook*, version 3.3, Ottawa.

Scharbatke-Church, C. 2011, 'Evaluating Peacebuilding: Not Yet All It Could Be' in Austin, B., Fischer, M. & Giessmann, H.J. eds, *Advancing Conflict Transformation: The Berghof Handbook II*, Opladen, Germany: Barbara Budrich: 460–482.

Shaw, R. 2005, *Rethinking Truth and Reconciliation Commissions: Lessons from Sierra Leone*, Washington: United States Institute of Peace.

Smith, D. 2004, 'Towards a Strategic Framework for Peacebuilding — Getting Their Act Together: Overview of the Joint Utstein Study of Peacebuilding', report commissioned by Royal Norwegian Ministry of Foreign Affairs, International Peace Research Institute, Oslo, April.

Smith, R.W. 2010, 'From Truth to Justice: How Does Amnesty Factor in? A Comparative Analysis of South Africa and Sierra Leone's Truth and Reconciliation Commissions', honors scholar thesis, University of Connecticut, Storrs.

Stover, H. & **Weinstein, H.M.** 2004 eds, 'Conclusions: A Common Objective, a Universe of Alternatives' in Stover, E. & Weinstein, H.M. eds, *My Neighbor, My Enemy: Justice and Community in the Aftermath of Mass Atrocity*, Cambridge, UK: Cambridge University Press: 323–342.

Tiemessen, A. 2008, 'Transitional Justice as a Normative Structure: What to Expect When You're Expecting Accountability', paper presented at Canadian Political Science Association Annual Conference, University of British Columbia, Vancouver, 4–6 June.

UN. 2010, 'Monitoring Peace Consolidation: United Nations Practitioners' Guide to Benchmarking', New York: http://www.un.org/en/peacebuilding/pbso/pdf/monitoring_peace_consolidation.pdf

UN. 2011, 'Strategic Results Framework on Women, Peace and Security: 2011–2020', New York, July: http://www.un.org/womenwatch/ianwge/taskforces/wps/Strategic_Framework_2011-2020.pdf

USAID. 2010, *Office of Inspector General: Audit of USAID/East Africa's Conflict Mitigation Activities*, Washington.

USAID. 2011a, 'Foreign Operations: FY 2011 Performance Report, FY 2013 Performance plan', Washington: http://www.usaid.gov/performance/apr/APR2011-2013.pdf

USAID. 2011b, *Mid-Term Evaluation of USAID's Counter-Extremism Programming in Africa*, Washington.

USAID. 2011c, 'Food for Peace Standard Indicators Handbook', Washington, December: http://transition.usaid.gov/our_work/humanitarian_assistance/ffp/ffpstdindicatorhb.pdf

Van Evera, S. 2003, 'Why States Believe Foolish Ideas: Nonself-Evaluation by States and Societies' in Hanami, A.K. eds, *Perspectives on Structural Realism*, New York: Palgrave Macmillan: 163–198.

Villa-Vicencio, C. 2009, *Walk with Us and Listen: Political Reconciliation in Africa*, Washington: Georgetown University Press.

From the margins to the mainstream: community restorative justice in northern ireland

Colin Knox

Published online: 19 Sep 2013.

PLEASE SCROLL DOWN FOR ARTICLE

FROM THE MARGINS TO THE MAINSTREAM: COMMUNITY RESTORATIVE JUSTICE IN NORTHERN IRELAND

COLIN KNOX

Abstract

Paramilitary organisations exerted a stranglehold on working class loyalist and republican communities in Northern Ireland during the conflict there. In the absence of an effective and legitimate policing service, paramilitaries developed an alternative 'justice' system in which they 'punished' those accused of committing crimes against the community. They adopted a punitive system of control which included threats or warnings, public humiliation, curfew, exiling, beatings and shootings. This article traces the evolution of this system from illegal paramilitary 'policing' through to restorative justice schemes offering non-violent alternatives to community crime, which, over time, have become a recognised part of the formal criminal justice system. Specifically, it examines the role that a series of evaluations had on influencing this transformation. At the very least, policy evaluation informed the political debate and provided evidence to move restorative justice from illegal activities to an integral part of the criminal justice system.

Keywords: restorative justice, paramilitaries, evaluation, Northern Ireland

Introduction

This paper traces how illegal practices of paramilitary groups involved in 'policing' their own communities in Northern Ireland shifted to adopt restorative justice schemes operating in the twilight of the law which were subsequently incorporated into the formal criminal justice system. Specifically, it will consider what role evaluations of these schemes played in community restorative justice moving from the margins to the mainstream of criminal justice policy. Three evaluations were conducted over a 10-year period: one by academics, one on behalf of funders (by Professor Harry Mika) and one through the independent statutory criminal justice inspectorate. The sequencing of the evaluations (see Table 1) coincided with significant political developments in Northern Ireland: the Belfast (Good Friday) Agreement 1998; periodic suspension and restoration of the devolved Assembly (2000 to 2007); republican endorsement of the policing (2007); and, finally, political stability and power-sharing (2007 onwards). Evaluation research had the potential to impact either positively or negatively on the wider peace process. Indeed, evidence gathered through the evaluations featured equally in support of how the political agreements reached were effective in embedding peace, and to prove that

57

Table 1: The Intersection of Evaluation and Political Context.

Evaluation of restorative justice in Northern Ireland	Political context
Funder: ESRC	• Fragile political talks leading up to the Belfast Agreement.
Dates: 1998–2000	• Involvement in paramilitary-style attacks could have excluded loyalist and republican parties from peace talks and political agreement
Methods: Interviews with victims, perpetrators, politicians, NGOs, police, probation board. Construction of database to understand nature and incidence of beatings and shootings	• British government adopts 'see no evil, hear no evil' approach to community violence — bigger prize stance
Funder: Atlantic Philanthropies (external funder of restorative justice interventions)	• Devolved government in place since 1999, although ongoing suspensions of the institutions
Dates: 1999–2005. Evaluation published January 2007	• St Andrews Agreement (2006) secures Sinn Féin support for policing and political deal on power-sharing arrangements
Methods: Extensive review of case files, large number of interviews with key stakeholders, and non-participant observation techniques	• Wider political imprimatur allows republican restorative justice schemes to work with police
Funder: CJI — non-departmental public body and independent statutory inspectorate funded from the public purse (formerly located in the Northern Ireland Office and now in the devolved Department of Justice)	• Much improved political milieu — power-sharing Executive and Assembly fully operational.
	• Active encouragement by CJI for restorative justice schemes to adopt protocol principles
Dates: 2007 onwards	• Restorative justice schemes refocus their work towards community mediation with a large number of statutory organisations
Methods: Evaluation inspections and follow-up reports using case files, organisational documentation and extensive interviews with key stakeholders	• Restorative justice mainstreamed and an integral part of formal criminal justice system.

some political parties were disingenuous about their long-term intentions. The paper is in three parts: first, the background to, and evolution of, the community informal 'justice' system and the transition into regulated restorative schemes; second, the detail of the evaluations that took place and provided an evidence base for this transition; and, third, the intersection of evaluation research and the 'high politics' of Northern Ireland.

Background

Northern Ireland has witnessed significant changes in its political, constitutional and security landscape since the beginning of the conflict there in 1969. Political stability has created an environment where violence is seen by all but the extremists as redundant. The existing political arrangements are rooted in the Belfast (Good Friday) Agreement 1998 that provided for, inter alia, a devolved Northern Ireland Assembly with full executive

and legislative authority, which has functioned continuously since 2007 after a series of faltering starts. Devolution wavered largely over the decommissioning of paramilitary weapons. From its inception in December 1999 until October 2002, the Assembly was suspended four times. A political breakthrough came in the form of the St Andrews Agreement in October 2006. Following elections, devolved power was restored to the Assembly in May 2007 with a power-sharing Executive headed by Ian Paisley as the Democratic Unionist Party (DUP) First Minister (now replaced by Peter Robinson) and Sinn Féin's Martin McGuinness as Deputy First Minister. A working system of local governance has been in place since 2007, described by the First Minister as the 'most settled period of devolution for over forty years' (Robinson 2009, 4). An uninterrupted period of devolution, the transfer of policing and justice powers to the Northern Ireland Assembly from Westminster (the so-called final piece of the 'devolution jigsaw') and a move away from constitutional and security issues herald a return to 'normal' politics.

Political violence in Northern Ireland cannot be completely discounted even though it has significantly decreased. There remains a small but significant level of violence perpetrated by paramilitary groups clinging to the vestiges of control in loyalist and republican working class areas. This community-based violence is the focus of this paper. Such was the role played by paramilitaries throughout the conflict in 'policing' their communities that an 'alternative criminal justice system' emerged with its own brutal punitive system of punishing wrong-doers or those accused of committing crimes against their own community.

> **Political violence in Northern Ireland cannot be completely discounted even though it has significantly decreased.**

The evolution of community informal 'justice'

Informal community 'justice' in contemporary Northern Ireland evolved in the early 1970s within Catholic working class communities, where citizen defence committees were set up to protect Catholic enclaves from loyalist attacks. As the security forces (the then Royal Ulster Constabulary [RUC]) withdrew from barricaded areas, local defence associations emerged to deal with petty crime within the community. Increasingly, paramilitary organisations became active in policing their own areas. Silke (1998, 124) described the spectrum of punitive actions or 'punishment scale' used. House or shop breakers were compelled to reimburse their victims and return stolen goods. In cases involving children, the Irish Republican Army (IRA) approached the parents and requested greater parental control. In situations where an alleged offender refused to cooperate or had ignored previous warnings, they were liable for 'suitable punishment'. This usually involved shooting or beating the individual anywhere in the leg — so-called 'kneecapping' (*Republican News* 1971, 2). For those 'too young to be kneecapped', 'punishments' included curfew, tar and feathers, being tied up and publicly painted, and reprimanding their parents (*An Phoblacht* 1982, 1). Individuals suspected of informing the police were dealt with most severely and their 'punishment' depended on the type of information passed to the security forces. In some cases they were kneecapped but usually they were shot dead. In loyalist areas, paramilitaries from the early 1970s assumed a similar policing role in their communities and used many of the methods adopted by republicans (Smithey 2011). Although paramilitary groups claimed to carry out investigations into incidents before an individual was punished, effectively they ignored due process, and the human rights of the accused were practically non-existent. Kennedy (1995, 14) described the system as a barbaric range of punitive measures against individuals 'who violated

some community norm as defined by the paramilitary grouping'. In short, the 1970s to mid 1990s period was characterised by wholly unofficial 'policing' of republican and loyalist communities by paramilitaries in de facto self-governing areas (Knox & Monaghan 2002).

Political progress and the paramilitary ceasefires in the mid 1990s caused a radical rethink of the informal 'justice' system. No longer could paramilitaries be involved in highly visible acts of violence, particularly against young people, and claim legitimacy in upholding ceasefires. As the political momentum accelerated in the form of the Belfast Agreement, pressure to loosen the grip of paramilitaries on communities increased. For example, the Belfast Agreement (1998, section 1: 4) outlined its 'opposition to any use or threat of force for any political purpose'. Alternative community 'policing' arrangements were therefore explored. In republican communities, the Northern Ireland Association for the Care and Rehabilitation of Offenders (NIACRO) approached Sinn Féin to investigate ways in which non-violent alternatives might be found to tackle community crime. The model that emerged was based on a system of community restorative justice developed in Canada and the US in the early 1970s. Community restorative justice has been defined as

> A more inclusive approach to dealing with the effects of the crime, which concentrates on restoring and repairing the relationship between the offender, the victim, and the community at large, and which typically includes reparative elements towards the victim and/or the community. (Criminal Justice Review Group, 2000, para. 1)

Loyalists developed a parallel scheme aimed at tackling anti-social behaviour. Both types of community restorative justice schemes, the loyalist Greater Shankill Alternatives Programme and the republican Community Restorative Justice Scheme (CRJI), received support funding from Atlantic Philanthropies, a US charitable foundation. These unregulated community restorative justice schemes operated from the late 1990s onwards with some success according to those involved in their operation (Auld et al. 1997; Winston & Watters 2006).

Regulating restorative justice

The unregulated community restorative system described above came under both general and legal pressures to adopt government regulation — put starkly, reform or be marginalised. The Belfast Agreement (1998, 22) argued that the police service must be 'capable of winning public confidence and acceptance, delivering a policing service in constructive and inclusive partnerships with the community at all levels, and with maximum delegation of authority and responsibility'. The follow-on political agreement at St Andrews (October 2006) reasserted the need for accountable policing by arguing, 'we have consistently said that support for policing and the rule of law should be extended to every part of the community. We believe that all parties share this objective' (Agreement at St Andrews 2006, section 5). The lack of confidence in policing, particularly in working class areas of Northern Ireland, had been the raison d'être for a parallel (rather than complementary) system of restorative justice. As confidence in the Police Service of Northern Ireland (PSNI) grew and the police gave their support to the principle of restorative justice, albeit with strict conditions, the pressure for change mounted.

There were also legal pressures on the government to regulate community restorative justice. The Human Rights Act 1998 was received with limited enthusiasm in Northern Ireland but the Belfast Agreement placed human rights at the centre of the political agenda. The Agreement went much further than the European Convention on Human Rights, recognising that Northern Ireland should be founded on the 'principles of full respect for, and equality of, civil, political, social and cultural rights', as well as 'freedom from discrimination for all citizens' (Belfast Agreement 1998, para. v). There followed two reports detailing how human rights should be implemented in practice. The Patten Report (1999) on policing reforms and the Criminal Justice Review (2000) were extensive documents recognising past institutional failings and recommending how human rights in Northern Ireland might be better protected in the future. Even though these 'additional' measures were not strictly legally binding obligations, the introduction of human-rights-friendly policies is both necessary and desirable to translate rhetorical respect into concrete observance (O'Cinneide 2006). However, it was these two reports (more so the latter) that really thrust community restorative justice onto the political agenda.

The protocol

As a result of the recommendations from the Criminal Justice Review, the (then) Minister of State for the Northern Ireland Office, David Hanson, published (after two attempts at consultation) the *Protocol for Community-Based Restorative Justice Schemes* (Northern Ireland Office 2007). In launching the protocol the Minister emphasised that it contained stringent safeguards to protect the rights of both victim and offenders, and that the police would be at the centre of the process. He claimed it

> put in place a structure which will provide for effective engagement between community-based schemes and the criminal justice system in dealing with low level offending. The high standards set out in the protocol are non-negotiable. (Hanson 2007, 2)

The protocol followed the precise headings of the Criminal Justice Review recommendations referred to above and included the following principles:

- Schemes must recognise that statutory responsibility for the investigation of crime rests with the police and that the only forum which can determine guilt or innocence, where this is at issue, is a court of law ...

- This means that any group or structures organised by the community should include provision for full co-operation and communication with the police. (Northern Ireland Office 2007, 3–4)

Not surprisingly, republican communities found this protocol totally unacceptable. The role of the police was central to reporting, investigating and applying sanctions at the community level. For republicans, at that time, this was simply unworkable. The police, on the other hand, argued that they needed to know the nature of the crime, who the offender was and that he or she was being dealt with by the scheme, since otherwise the offender would be left open to double jeopardy. Acutely aware, however, of the need to satisfy standards by which its schemes should operate, republicans (drawing on previous research) developed their own code of practice, setting standards pertaining to participants and the community, and outlining fundamental concepts of restorative justice, without reference to the police in their documentation (Community Restorative Justice 1999). The Northern Ireland Office protocol included an accreditation process

whereby each community-based restorative justice scheme had to confirm to the Criminal Justice Inspection (CJI), an independent statutory body with responsibility for inspecting all aspects of the criminal justice system in Northern Ireland (apart from the judiciary), its willingness to adhere to the protocol. If, after inspection, the inspectorate was satisfied that the standards set out in the protocol were being met, the scheme became accredited by the Northern Ireland Office. Schemes that received accreditation became eligible for government resources (the carrot); those schemes which did not apply or were turned down did not receive government funding or engage formally with the criminal justice system (the stick). The latter could continue with their unregulated work and, providing they did nothing illegal, the government was in no position to discontinue their work.

> **Schemes that received accreditation became eligible for government resources (the carrot); those schemes which did not apply or were turned down did not receive government funding or engage formally with the criminal justice system (the stick).**

From the margins to the mainstream

The loyalist Northern Ireland Alternatives (NIA) and its four restorative schemes were the subject of an inspection with a view to accreditation from the CJI in April 2007. The report noted the role played by the schemes in reducing 'punishment' beatings by offering alternative means of dealing with low-level offending: 'the schemes … first and foremost are a community resource dedicated to working with difficult youngsters, either diverting them away from crime in the first place or helping them to draw them out of criminal and anti-social behaviour' (CJI 2007a, section 3.6, 18).

The report concluded that the loyalist schemes 'worked to a high standard with difficult young people in their communities' and recommended that NIA be accredited, subject to agreement on conditions identified by the inspectorate team. The first schemes to be accredited were: the loyalist NIA, Greater Shankill Alternatives, East Belfast Alternatives and North Belfast Alternatives.

In October 2007, the CJI completed a pre-inspection report of republican schemes in Belfast and Derry/Londonderry which operated under the auspices of CRJI, the first stage of the accreditation process. Their involvement in this initial process was significant and was helped by the fact that in January 2007 Sinn Féin publicly recognised the PSNI. The report covered two sets of CRJI schemes — four in Derry and four in West Belfast and essentially described the state of readiness of the schemes for accreditation. The inspectors concluded, the fact that, for historical reasons, the schemes do not normally pass information to the police means that they are not at present operating in accordance with the Protocol' (CJI 2007b, section 5.2, 31). The report recommended that the CRJI schemes should be considered for accreditation as soon as they were ready to declare that they were complying with the protocol, and set out several suggestions as to how they might do this.

In June 2008 the CJI re-inspected the republican restorative justice schemes and, based on a balance between 'risks and opportunities involved', they confirmed accreditation (CJI 2008, 13). The risks were of the schemes behaving improperly, and the opportunities were establishing a proper relationship with the police and helping them to reach out to communities that were alienated and poorly served. In accrediting the schemes the inspectorate noted that they needed to be closely monitored and that a fully independent complaints mechanism should put in place. In a follow-up report conducted by the

inspectorate in 2011, CRJI was criticised for their complaints policy and not meeting the threshold for inclusion under the government's protocol for community-based restorative justice. However, the report concluded that

> The schemes are increasingly assuming a role where they are attempting to influence dissident republican paramilitaries and other armed groups away from intimidation, expulsions and violence. This work is valued particularly by the police, local community and some political leadership and is seen as an important element in ensuring that these events do not become commonplace. (CJI 2011, 19, 5.7)

The schemes are involved in the delivery of services ranging from community and mediation support, counselling and youth work, to interventions in cases of anti-social behaviour and those 'under threat'. The schemes receive financial support from a range of statutory and philanthropic funders, all of which express a high level of satisfaction with the services they provide (Knox 2011). The statutory organisations include: the Department of Justice, Department for Social Development, Probation Board for Northern Ireland, PSNI, Northern Ireland Housing Executive and the Belfast Health Trust.

The transformative nature of the fully accredited restorative justice schemes operated through CRJI and NIA is now recognised by most statutory agencies on the ground.

The key stages in the transition of restorative justice schemes from the margins to the mainstream are summarised in Table 2.

Evaluation of the Schemes

Given the journey of the restorative justice schemes, from operating as illegal activities directed by paramilitaries to becoming an integral element in the formal criminal justice system, a key question in this paper is the extent to which formal evaluation and evidence therein played a part in this process. Evaluation research, according to Rossi and Freeman (1993, 5), is 'the systematic application of social research procedures for assessing the conceptualisation, design, implementation and utility of social intervention programs'. Three evaluations were conducted during the transition from illegal practices to mainstream restorative justice; they were very different in focus and type. The first evaluation, using taxonomy developed by Patton, might be described as a 'responsive evaluation' which sought to capture, represent and interpret varying perspectives, in particular why communities 'accepted' domination by paramilitary groups and why government was prepared to turn a blind eye to this practice (Patton 2008; Stake & Abma 2005). The second evaluation can be understood as an 'impact evaluation' which considered the outcomes of community restorative models for victims and perpetrators and whether this type of intervention was an effective way of tackling low-level crime and anti-social behaviour. The third evaluation is depicted as having a 'compliance focus': were restorative justice schemes acting in accordance with human rights principles and criminal justice protocol arrangements? We now consider each of these in some detail.

Table 2: Community-Based Restorative Justice Schemes — the Transition.

Informal 'justice' (1970 to mid 1990s)	Unregulated community restorative justice (1994–2006)	Regulated community restorative justice (2007 onwards)
• Absence of legitimate policing service in loyalist and republican areas	• Non-violent alternative schemes set up: NIA (loyalist) and CRJI (republican)	• Government produces protocol (2007) to accredit community-based schemes
• Paramilitary organisations 'police' their own communities	• Schemes operate outside government/police control	• New regulated schemes can deal only with low-level offences referred to them by Public Prosecution Service
• Tariff system of 'punishment' operates	• Loyalists 'cooperate' with police; republicans eschew RUC/PSNI	• Schemes required to operate in accordance with the Human Rights Act 1998 and UN Convention on the Rights of the Child in their interaction with young victims and offenders
• Due process ignored and no protection for rights of alleged perpetrators	• Schemes funded by philanthropic sources	• NIA and CRJI receive accreditation (2007 and 2008, respectively) from CJI
		• Schemes provide restorative services to a range of statutory organisations

Academic research (responsive evaluation)

The first project, entitled 'An Evaluation of the Alternative Criminal Justice System in Northern Ireland', was conducted by academics and funded by the respected and independent Economic and Social Research Council (ESRC) in the UK. The overall aim of the research was to contribute to an understanding of increasing paramilitary influence in 'policing' the two communities in Northern Ireland and the concomitant threat to social stability. The research took place from 1998 to 2000 and involved significant primary data collection — 42 interviews with victims, two with perpetrators, and six focus groups with key stakeholders including the police and probation services, 12 politicians and 80 statutory and voluntary/non-governmental organisations in Northern Ireland. In addition, two large Excel databases of reported paramilitary-style attacks were constructed and geographic information system (GIS) maps of locations were drawn.

The timing of the evaluation is significant because it straddled important political events. The Belfast Agreement was signed on 10 April 1998, following which power was devolved from Westminster to the Northern Ireland Assembly on 2 December 1999. During the negotiations in the run-up to the Agreement all parties involved had to endorse the Mitchell principles of democracy and non-violence. One principle urged that 'punishment' killings and beatings stop and parties take effective steps to prevent such actions. The negotiators noted, 'We join the governments, religious leaders and many others in condemning "punishment" killings and beatings' (Mitchell et al. 1996, para. 20).

The outcome of the multi-party talks was the Belfast Agreement, in which all participants reaffirmed their 'total and absolute commitment to exclusively democratic and peaceful means of resolving differences on political issues, and our opposition to any threat of force by others for any political purpose' (Belfast Agreement 1998, section 4:1). Over 19 months later, in advance of devolution, Sinn Féin stated the importance of the political process in making conflict a thing of the past and emphasised their opposition to the use of force and 'punishment' attacks (Sinn Féin statement, 16 November 1999).

The findings of the evaluation were damning in the extreme. The researchers found that victims of paramilitary 'punishment' beatings and shootings had become *expendable* and *legitimate* targets for violence in Northern Ireland. They were expendable in the sense that any attempt to deal with this problem in a serious way would have widespread political ramifications for parties currently in devolved government. The Mitchell principles of 'democracy and non-violence' which were pivotal to the Belfast Agreement could not be upheld. It appears that it was political expedient to turn a blind eye to these acts of brutality. They were legitimate in the sense that the victims' culpability derived from the communities within which they lived and their 'punishment' was meted out by paramilitaries acting on the communities' behalf. Both these factors conspired to make this group not only the forgotten victims of violence in Northern Ireland, but arguably the most vulnerable. The researchers described this as a 'see no evil, hear no evil' reaction on the part of the British government given the fragility of the peace process (Knox 2002). The government's response was that one must accept certain violent excesses in the interests of moving forward politically. This justification appears in the lexicon of political debate on Northern Ireland at the time. The Secretary of State, Mo Mowlam (1999, 2), suggested, 'the peace we have now is imperfect, but better than none' or perhaps, more tersely, as having 'an acceptable level of violence'.

The research also became the subject of two parliamentary debates. In the House of Commons the Conservative Party attempted to halt the early release of political prisoners under the terms of the Belfast Agreement because of continuing paramilitary-style shootings and beatings (Hansard, Official Report, 1999). In a Northern Ireland Assembly debate, pro- and anti-Agreement politicians adopted contrasting positions on the ESRC evaluation findings. One pro-Agreement party argued that the report 'highlighted the strong support that there is for alternatives to "punishment" attacks (such as restorative justice) in the absence of a legitimate policing service' (Gildernew 2001, 361). Anti-Agreement parties claimed that paramilitary attacks had soared as a direct result of prisoner releases and the research 'presented cogent evidence that the Good Friday Agreement is failing' (Paisley 2001, 366). Politicisation was an ongoing feature of the debate on the issue. The (then) Sinn Féin Health Minister, Bairbre de Brun, was regularly asked for information in the Northern Ireland Assembly on how the immediate hospitalisation of those subject to beatings and shootings was displacing patients on long waiting lists in need of orthopaedic surgery and trauma counselling (Hansard, Official Report, 2001). This was as much intended to embarrass and undermine the Sinn Féin minister as it was to highlight the plight of paramilitary victims.

The funder's evaluation (impact evaluation)

The second evaluation was commissioned by the external funder (Atlantic Philanthropies) of the restorative justice schemes, NIA and CJRI. It was conducted by Professor Harry Mika, an international expert on restorative justice, from Central Michigan University and the Institute of Criminology and Criminal Justice in Queen's University,

Belfast. Mika's primary research involved the analysis of 500 formal case interventions and 'several hundred interviews' over the period between 1999 and 2005.

The findings of the evaluation concluded, inter alia, the following:

- The restorative justice programmes prevented nearly 500 cases of paramilitary beatings and shootings. NIA and CRJI caused a significant drop in the number of beatings and shootings compared to neighbourhood areas outside their catchment population.

- The acceptance of community restorative justice solutions by armed groups increased significantly.

- Community leaders felt that the projects had become essential community assets.

- Potential limitations of NIA and CRJI were caused by: perceived paramilitary links, political criticism, inadequate resourcing and increasing demands for their services. (Mika 2007, 33–34)

Mika's evaluation was a very strong endorsement of the community restorative justice schemes and also coincided with a series of political events that created the context for a more receptive response than the ESRC evaluation. The Northern Ireland Assembly had been faltering since its inception in 1999 and was dissolved by the British government in October 2002. The two issues of central concern which formed the basis of talks to achieve restoration of the political institutions were: the need to support policing and the rule of law across the whole community and eventually the devolution of policing and justice, and support for power-sharing and the political institutions. The British and Irish governments reached agreement on these issues as set out in the St Andrews Agreement of 13 October 2006, the details of which were given legislative effect in the Northern Ireland (St Andrews Agreement) Act 2006. The Act made provisions for a new transitional Assembly, set out a timetable to restore devolution, including the date for the third election to the Northern Ireland Assembly, and made important amendments to the Northern Ireland Act 1998 which came into force with the restoration of devolved government on 8 May 2007. Ian Paisley (then DUP leader) and Gerry Adams (the leader of Sinn Féin) agreed to establish a power-sharing Executive which has been in operation since then. Very quickly thereafter Sinn Féin joined the Northern Ireland Policing Board (June 2007) for the first time since it was established in 2001. The role of the board is to hold the PSNI to account through the chief constable for the delivery of effective and impartial policing.

Sinn Féin's members had backed their leadership's proposal to get involved in policing on the condition that a power-sharing Executive was established. With power-sharing in place and Sinn Féin participating in the Policing Board, there was no reason why restorative justice schemes in republican areas should not be cooperating fully with the PSNI, one of the key stumbling blocks to the accreditation process. The political choreography continued when CRJI wrote to the Minister of State for Northern Ireland in February 2007 seeking accreditation of their restorative justice schemes in Belfast and Derry/Londonderry. In short, Mika's evaluation findings coincided precisely with a time when political developments were wholly positive towards a response to republican overtures on policing and security, including their role in restorative justice schemes. The accreditation process by the CJI was to copper-fasten republican commitments to community restorative justice.

The accreditation process by the CJI was to copper-fasten republican commitments to community restorative justice.

Criminal Justice Inspection reports (compliance evaluation)

The inspection reports of the CJI were key documents in moving restorative justice schemes from the margins to the mainstream of the criminal justice system. The loyalist NIA schemes were quick to avail themselves of the opportunity to achieve formal accreditation given their ongoing cooperation with the PSNI and sought an inspection in March 2007. The methodology used by the inspectors was to gather both primary and secondary research data. They read extensive documentation: case files, policy documents, training materials, management committee minutes and annual reports and accounts. They also interviewed staff and volunteers, clients of the schemes (both young offenders and victims), parents of the children participating in the schemes and a wide range of other interested parties, including local politicians, PSNI officers, probation officers and school teachers with experience of working with the schemes.

Fears that community-based restorative justice schemes were a front for paramilitary organisations or that people were forced into taking part in restorative justice by paramilitaries were addressed in the course of the inspections by the CJI. The inspectors found no evidence that there was any such problem in relation to NIA or its schemes. In addition, there was no evidence of the schemes being driven by paramilitaries and every indication to the contrary. The report concluded that Alternatives 'did not provide an alternative policing or judicial system. Most of the work undertaken by the schemes relates to community development' (CJI 2007a). The inspectors supported accreditation. A follow-up report was conducted by the CJI in February 2010 and endorsed the earlier positive evaluation. The inspectors concluded that they 'had heard unanimous support for the work of NIA and the contribution the organisation was making in helping the lives of people living within some of the most socially deprived loyalist areas of Greater Belfast and North Down'. As a result of NIA's willingness to engage with statutory agencies they had earned 'real respect and a desire to increase the level of partnership working' (CJI 2010, 14).

Republican restorative schemes struggled with the accreditation process and failed to achieve the same recognition as NIA. A pre-inspection report conducted by the CJI in May 2007 found that although the republican schemes were engaged in work that was valued by their communities, there were improvements that needed to be made before the schemes would be ready for accreditation (e.g. training staff to work to the standard required by the protocol and to improve their record-keeping and the secure storage of files). A further inspection took place in June 2008 to determine CJRI's accreditation status. The inspectors found that the schemes were operating lawfully and non-coercively, were respecting human rights and were beginning to develop a constructive relationship with the PSNI. All 10 schemes operated by CRJI were accredited as a result (July 2008). However, in a follow-up inspection involving an examination of case files, CJI found that since the securing of accreditation only one case had been referred by CRJI to the PSNI under the government protocol, which highlighted a need for the current protocol to be reviewed. Despite this, the inspectors found a number of positive developments had occurred in the three years since its previous inspection. 'CRJI has become an important part of the voluntary and community sector landscape in parts of Northern Ireland and are integrating their activities as part of local community safety networks' (CJI 2011, 13).

Conclusions

It is clear that the political context in Northern Ireland has improved significantly and, with that, restorative justice schemes have moved from their original mission of providing an alternative option for young people who 'came to the attention' of paramilitaries because of anti-social behaviour. Schemes now work in partnership with many statutory organisations, using their restorative practice skills in other public policy areas where mediation is required (e.g. education and neighbourhood disputes). What is also clear is that as political stability has become embedded, mutilations, torture, beatings and exiling can no longer come within the purview of an 'acceptable level of violence' or be seen as part of the imperfections of the peace process. The inspections/evaluations conducted by the CJI have played an important part in 'normalising' restorative justice schemes. The 'risks and opportunities' calculus they use — schemes breaching human rights of offenders versus the potential for republicans and loyalists to work in partnership with the police — has paid off. All three evaluations demonstrate the inextricable link between the changing political context and efforts to mainstream community restorative justice. Evidence from the ESRC evaluation became part of the vitriolic pro- and anti-Agreement debate. Mika's evaluation coincided with a period of political consensus on policing and power-sharing. And the CJI reports ushered restorative justice schemes from the margins into mainstream criminal justice policy. However, the direction of the influence is unclear here. Did the evaluations offer valuable empirical evidence at critical junctures in the political process, or did improving peace-building efforts allow a much more conducive environment in which restorative schemes could flourish?

> **What is also clear is that as political stability has become embedded, mutilations, torture, beatings and exiling can no longer come within the purview of an 'acceptable level of violence' or be seen as part of the imperfections of the peace process.**

There is also the question of whether the nature of the evaluations was important to the influence they had? The ESRC study or responsive evaluation (to return to Patton's taxonomy above) conducted by academics presented independent evidence regardless of its impact on the fragile political environment. This evaluation offered an in-depth analysis of key stakeholders in the process — paramilitaries, police, victims and government. It challenged the Northern Ireland Secretary of State to confront the contradictions of an 'acceptable level of violence'. The potential impact of this evaluation was limited by the politicisation of its findings but it provided important evidence as to why communities 'accepted' paramilitaries as guarantors of local justice. The funders' or impact evaluation (in Patton's terms) was straightforwardly an attempt to assess whether externally supported interventions were an effective alternative way of dealing with a repressive 'justice' regime. The strength of this study was the duration of the research which allowed evaluators to capture change over time and affirm NGOs as recognised mediators in the restorative process. Its findings were timely in that they coincided with a period when working with the police became consistent with the wider republican political agenda. In other words, there is evidence that restorative justice works and the question became how best to mainstream an externally funded intervention. The CJI reports or compliance evaluation (in Patton's taxonomy) became a passport to government funding and legitimacy with other statutory organisations because they offered evidence of restorative justice schemes complying with international human rights standards that respected due legal process and the rights of the victims and perpetrators. Given the source of these reports and the weight they carried as a

consequence, it seems likely that they were pivotal to the success of incorporating restorative justice as part of the formal criminal justice system. Collectively, therefore, evaluation research shone a light on the illegal activities of paramilitaries in 'policing' their communities and the acquiescence of the police and government in this process; provided evidence of an effective alternative administered through local NGOs with experience in restorative justice; and demonstrated how restorative schemes could provide services that were fully compliant with the law and human rights standards. The three evaluations of restorative justice schemes in Northern Ireland and their intersection with the 'high politics' of Northern Ireland are summarised in Table 1.

Is it possible to be definitive about the impact of restorative justice schemes? If one looks at the statistics since 1982 categorised as paramilitary-style shootings and beatings, some patterns emerge. Figure 1 shows the combined figures for beatings and shootings by republican and loyalist paramilitaries over time (PSNI 2012).

The paramilitary ceasefires of August and October 1994 witnessed a significant decrease in the number of *shootings* to the lowest recorded level, but *beatings* simultaneously increased to their highest recorded level, hence the spike in Figure 1. This was a technical cop-out by the paramilitaries, who could claim they were not breaking the conditions of the ceasefires — instead of shooting those involved in anti-social behaviour, they beat them. The introduction of restorative justice programmes seemed to have had a short-term effect on the overall level of paramilitary attacks but numbers increased to the

The introduction of restorative justice programmes seemed to have had a short-term effect on the overall level of paramilitary attacks but numbers increased to the highest recorded levels in 2001.

highest recorded levels in 2001. During that period, wider political developments in the peace process were in trouble. The Northern Ireland Assembly was indefinitely suspended in October 2002 for the fourth time since devolution (December 1999) due to 'a lack of trust and loss of confidence on both sides of the community', according to the Secretary of State (Reid 2002, 201). This stemmed from concerns about Sinn Féin's commitment to exclusively democratic and non-violent means and accusations by each community of the other that they did not endorse the full operation and implementation of the Belfast Agreement. A climate of mistrust and uncertainty prevailed, accentuated by events such as the trial of republicans in Colombia (allegedly involved in training the left-wing Revolutionary Armed Forces of Colombia [FARC]), but subsequently found not

Figure 1: Paramilitary-Style Attacks.
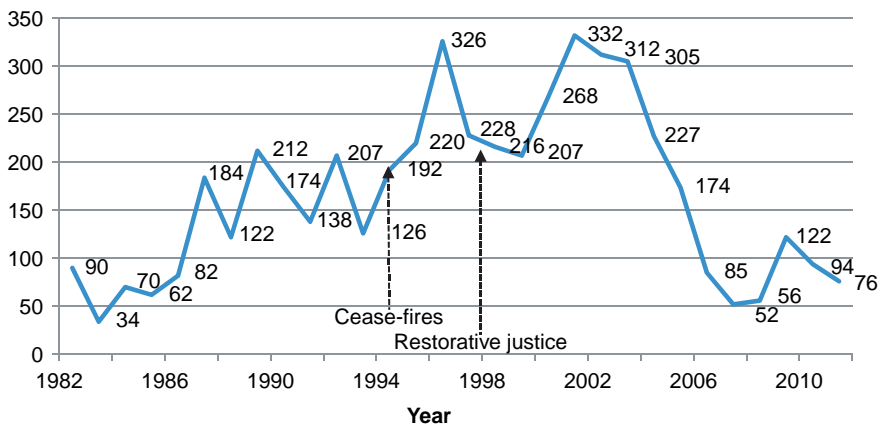
guilty), the break-in at Special Branch offices in Castlereagh Police Station (where personal details of Special Branch detectives were removed) and political espionage at Stormont implicating Sinn Féin. In the latter, unionists accused the IRA of exploiting Sinn Féin's membership of the Executive to gather information (names and addresses of prison officers) of use in future acts of violence. Since 2002, the trend in paramilitary-style attacks has been downward although the problem has not been eradicated because loyalist paramilitaries now involved in drug-dealing do resort to their old punitive tactics.

Have the restorative justice schemes contributed to the downward trend in paramilitary-style attacks? There are limitations to assessing impact through examining statistics, not least because of our inability to isolate cause and effect variables in the restorative justice schemes and. In addition, there is the problem of establishing the counterfactual position; in the absence of the schemes, could beatings and shootings have been a lot higher? The level of 'punishment' attacks may have little to do with what happens within the restorative schemes and be influenced by extraneous factors over which they have no control. The loyalist turf feud between factional paramilitary groups in the Shankill (area of Belfast) in 2000, for example, probably resulted in a number of 'housekeeping' attacks by paramilitaries. Developments in the wider political landscape such as police reforms and the changes in the criminal justice system have impacted on the continuance or otherwise of 'punishment' attacks. It is also unrealistic to expect restorative justice schemes in isolation to tackle the systemic causes of anti-social behaviour: poverty, unemployment, urban decay and the wider social, political and economic milieu in which community violence exists. Disen-

> Policy evaluations of the schemes informed the political debate and, by design or default, provided timely evidence that helped to move restorative justice from illegal activities that paramilitaries used to exert control in working class communities to an integral part of the criminal justice system.

tangling the evidence emerging from evaluations of restorative justice schemes in a context of significant political reform is problematic. Proving a cause and effect relationship in the pattern of declining paramilitary-style attacks is equally difficult. At the very least, policy evaluations of the schemes informed the political debate and, by design or default, provided timely evidence that helped to move restorative justice from illegal activities that paramilitaries used to exert control in working class communities to an integral part of the criminal justice system.

COLIN KNOX is Professor of Comparative Public Policy in the School of Criminology, Politics and Social Policy at the University of Ulster, Northern Ireland.

References

Agreement at St Andrews 2006, Belfast: Northern Ireland Office.

An Phoblacht 1982, 'Punishment Shootings', 25 March: 1.

Auld, J., **Gormally, B.**, **McEvoy, K.** & **Ritchie, M.** 1997, *Designing a System of Restorative Justice in Northern Ireland: A Discussion Document*, Belfast: Institute of Criminology and Criminal Justice, Queen's University.

Belfast Agreement 1998, *Agreement Reached in Multi-Party Negotiations*, Belfast: Northern Ireland Office.

CJI 2007a, *Northern Ireland Alternatives: Report of an Inspection with a View to Accreditation under the Government's Protocol for Community Based Restorative Justice*, Belfast.

CJI 2007b, *Community Restorative Justice Ireland: Report of a Pre-inspection of Schemes in Belfast and in the North West with a View to Accreditation under the Government's Protocol for Community Based Restorative Justice*, Belfast.

CJI 2008, *Community Restorative Justice Ireland: Report of an Inspection*, Belfast.

CJI 2010, *A Follow-up Review of the Community Restorative Justice Schemes Operated by Northern Ireland Alternatives*, Belfast.

CJI 2011, *Community Restorative Justice Ireland: A Follow-up Review of the Community Restorative Justice Ireland Community Restorative Justice Schemes*, Belfast.

Community Restorative Justice 1999, *Standards and Values of Community Restorative Justice*, Belfast.

Criminal Justice Review Group 2000, *Review of the Criminal Justice System in Northern Ireland*, London: Stationery Office.

Gildernew, M. 2001, 'Northern Ireland Assembly Debate' in *Punishment Beatings, Hansard, Official Report*, 23 January: 356–370.

Hansard, Official Report 1999, 'House of Commons Debate' in *Terrorist Mutilations (Northern Ireland)*, 27 January: 347–398.

Hansard, Official Report 2001, 'Northern Ireland Assembly Debate' in *Paramilitary Activity and Hospital Waiting Lists*, 18 September: 201–240.

Hanson, D. 2007, 'Stringent Safeguards in New CRJ Protocols' in *News Release, Northern Ireland Office*, 5 February.

Kennedy, L. 1995, 'Nightmares within Nightmares: Paramilitary Repression within Working-Class Communities' in Kennedy, L. eds, *Crime and Punishment in West Belfast*, Belfast: Summer School, West Belfast: 22–35.

Knox, C. 2002, '"See No Evil, Hear No Evil": Insidious paramilitary violence in Northern Ireland' in *British Journal of Criminology* 42: 1: 164–185.

Knox, C. 2011, *Community Based Restorative Practice Schemes: A Review of Funding Arrangements*, Belfast: Atlantic Philanthropies.

Knox, C. & **Monaghan, R.** 2002, *Informal Justice in Divided Societies: Northern Ireland and South Africa*, Basingstoke: Palgrave Macmillan.

Mika, H. 2007, *Community-Based Restorative Justice in Northern Ireland*, Belfast: Queen's University.

Mitchell, G., **de Chastelain, J.** & **Holkeri, H.** 1996, *Report of the International Body on Arms Decommissioning*, Belfast: Northern Ireland Office.

Mowlam, M. 1999, *Secretary of State's Judgement on Alleged Breach of Cease-Fire*, Belfast: Northern Ireland Office.

Northern Ireland Office 2007, *Protocol for Community-Based Restorative Justice Schemes*, Belfast.

O'Cinneide, C. 2006, 'Equivalence in Promoting Equality', www.equality.ie

Paisley, I. 2001, 'Northern Ireland Assembly Debate' in *Punishment Beatings, Hansard, Official Report* 23 January: 356–370.

Patten Report 1999, *A New Beginning: Policing in Northern Ireland: Report of the Independent Commission on Policing in Northern Ireland*, Belfast: Stationery Office.

Patton, M.Q. 2008, *Utilization-Focused Evaluation*, London: Sage, 4th edn.

PSNI 2012, Police Service of Northern Ireland Statistics: http://www.psni.police.uk/index/updates/updates_statistics.htm, accessed 15 October 2012.

Reid, J. 2002, Statement by the Secretary of State for Northern Ireland to the House of Commons, Hansard, Official Report, 15 October.

Republican News 1971, *An Phoblacht*, 'Why Tar and Feathers?' 20 February: 2.

Robinson, P. 2009, 'Making Devolution Work', speech by Peter Robinson at Ulster Hall, Belfast, 8 September.

Rossi, P.H. & **Freeman, H.E.** 1993, *Evaluation: A Systematic Approach*, London: Sage, 5th edn.

Silke, A. 1998, 'The Lords of Discipline: The Methods and Motives of Paramilitary Vigilantism in Northern Ireland' in in *Low Intensity Conflict and Law Enforcement* 7: 2: 121–156.

Smithey, L. 2011, *Unionists, Loyalists, and Conflict Transformation in Northern Ireland*, Oxford: Oxford University Press.

Stake, R. & **Abma, T.** 2005, 'Responsive Evaluation' in Mathison, S. eds, *Encyclopaedia of Evaluation*, London: Sage: 376–379.

Winston, T. & **Watters, D.** 2006, 'Evidence given to the Northern Ireland Affairs Committee' in *Draft Protocol for Community-based Restorative Justice Schemes* Hansard, Official Report HC 87: 37–46.

Landmines and Livelihoods in Afghanistan: Evaluating the Benefits of Mine Action

Ted Paterson , Barry Pound & Abdul Qudous Ziaee
Published online: 19 Sep 2013.

PLEASE SCROLL DOWN FOR ARTICLE

LANDMINES AND LIVELIHOODS IN AFGHANISTAN: EVALUATING THE BENEFITS OF MINE ACTION

TED PATERSON, BARRY POUND AND ABDUL QUDOUS ZIAEE

Abstract

Mine action started in Afghanistan and, globally, has grown into a billion-dollar endeavour. On most measures, Afghanistan remains the world's largest mine action programme, which has performed admirably in terms of delivering outputs such as square metres cleared and devices destroyed. But less is known about when and how mine action enhances the well-being of people in mine-affected communities. This article outlines how the Sustainable Livelihoods approach has been used to provide a better understanding of the benefits of mine action, and how capacities have been developed to conduct future evaluations without dependence on international specialists, reducing both costs and risks.

Keywords: mine action, landmines, Afghanistan, Sustainable Livelihoods, evaluation, evaluation capacity development

Introduction

Afghanistan epitomises modern conceptions of a violently divided society and serves up an array of challenges for evaluators. This article describes how some international and Afghan organisations employed the Sustainable Livelihoods (SL) approach to evaluate the outcomes stemming from 'humanitarian demining'[1] and how evaluators coped with the security and safety concerns of both the evaluation teams and the residents of communities visited, as well as power imbalances (in particular, along gender lines), competing accountabilities, and risks of unintentional harm.

Beginning with a brief tour of the mine action sector, the article turns to evaluation capacity development in the Mine Action Program for Afghanistan (MAPA), the design and conduct of two Landmines & Livelihoods (L&L) surveys, some of the principal findings of these surveys,[2] and the subsequent steps taken by the Mine Action Coordination Centre for Afghanistan (MACCA)[3] to implement recommendations. The authors suggest that fostering local evaluation capacity, in addition to its intrinsic desirability, provides a partial answer to the question of how to overcome evaluation challenges in violently divided societies.

73

Mine Action and Evaluation Practices: Globally and in Afghanistan

Mine action aims to eliminate the threat posed by landmines and unexploded ordnance (UXO) on civilians. In heavily contaminated areas, demining is a necessary precondition for the safe delivery of humanitarian, peacebuilding, and development services, and for people's own efforts to rebuild their farms, businesses, and communities. Aid-financed mine action is a comparatively new endeavour, beginning in Afghanistan shortly Soviet forces withdrew in 1989. Global donor funding has averaged US$455 million per year over the past five years, with Afghanistan the world's largest recipient, accounting for 22% of the global total[4] (International Campaign to Ban Landmines [ICBL] 2012a, 49). Since 1991, MACCA has tallied US$1.05 billion in aid financing for mine action (Government of the Islamic Republic of Afghanistan [GIRA] 2012, 145).

> **In heavily contaminated areas, demining is a necessary precondition for the safe delivery of humanitarian, peacebuilding, and development services, and for people's own efforts to rebuild their farms, businesses, and communities.**

Mine action, and particularly demining, often receives significant funding when the operating environment is parlous because operators are organised with paramilitary discipline, because it has a good 'burn rate', helping donors meet well-publicised spending commitments, and because many armed groups support demining (in Afghanistan, Mullah Omar, the spiritual leader of the Taliban, wrote a letter enjoining supporters to facilitate mine action because demining was a form of jihad). During the Taliban regime, funding for mine action represented up to 13% of official development assistance to Afghanistan.[5]

Since 1989, over 1,210 square kilometres of suspected hazardous areas (SHA) have been demined (GIRA 2012, 6). Estimated mine/UXO casualties fell from over 9,000 per year in the mid-1990s to just over 800 in 2011 (ICBL 2012a, 33). In spite of this progress, 5,660 SHA remain, affecting 617 square kilometres and over one million people: eliminating this will take another decade and cost US$615 million (ICBL 2012b, 7).

Afghanistan has also featured in the international community's efforts to evaluate mine action. Initially, monitoring and evaluation (M&E) contributed only modestly to understanding the developmental benefits stemming from mine action (Paterson 2005). From its early days, the mine action sector has been adept in generating comforting figures on outputs: square metres cleared, landmines destroyed, etc. Output measures are necessary but they provide little insight into when and how mine action enhances the well-being of people in mine-affected communities.

Early evaluations focused on activities financed by individual donors, meeting the donor's accountability requirements perhaps, but failing to provide much understanding of how a national mine action programme contributes to development and to saving lives. Some improvements began to emerge in 1997, when the UN commissioned a programme-wide evaluation of mine risk education in Afghanistan (Andersson et al. 1998). In 2000, Britain, Canada, and Japan commissioned a joint evaluation of the MAPA (Van Ree et al. 2001). About the same time, two socio-economic assessments of the demining programme were released by the World Bank and the Mine Clearance Planning Agency (MCPA — an Afghan NGO).

Broadly, the evaluations were positive. In particular, the MCPA study (*Socio-Economic Impact Study of Landmines and Mine Action Operations in Afghanistan* [SEIS]) and that of the

World Bank (*Socio-economic Impact of Mine Action in Afghanistan* [SIMAA] [Byrd and Glidestad 2001]) arrived at similar estimates of the economic benefits of demining — roughly US$30 million in benefits for US$25 million spent. These similar conclusions suggested the estimates were reliable in spite of the difficulties in obtaining economic data in a conflict-affected country.

However, scratch the surface and key similarities disappear. As depicted in Figure 1, the SEIS concluded that 70% of total benefits stemmed from livestock production (largely from grazing land), with crops accounting for less than 5% of the benefits. The SIMAA concluded the exact opposite, with almost two-thirds of all benefits coming from crops and under 5% from livestock.

While some of these variances probably stem from differences in the classification of benefits (e.g. livestock in the SEIS compared with grazing land in the SIMAA), and in analytic methods (e.g. MCPA did not use discounting in its cost–benefit calculations), it seems something more fundamental was at play. One hypothesis was that the MCPA and World Bank teams had different understandings of the nature of the rural economy, and looked for different things. The Bank's analysis seemed grounded in a perspective that crop production is the main factor in rural livelihoods, whereas research at the time suggested a more complex pattern in Afghanistan (Christoplos 2004). For example, grain production and livestock typically form an integrated farming system, with dung used as a fertiliser and the crop chaff as animal fodder (Maletta 2004), and economic analysis should be based on the farming system rather than individual components.

Such questions are critical in Afghanistan and other conflict-affected countries because it takes decades to find and clear the mines and UXO from rural areas. Should priorities be set on the basis of a cost–benefit calculus, or is a more community-based approach required to understand how demining can best contribute to sustainable livelihoods? The former approach promises greater efficiency and cost-effectiveness, as well as being readily monitored and in accord with the command-and-control management styles that the humanitarian demining operators employ.

Figure 1: Findings from the SEIS and SIMAA (Paterson 2005, 318–319)
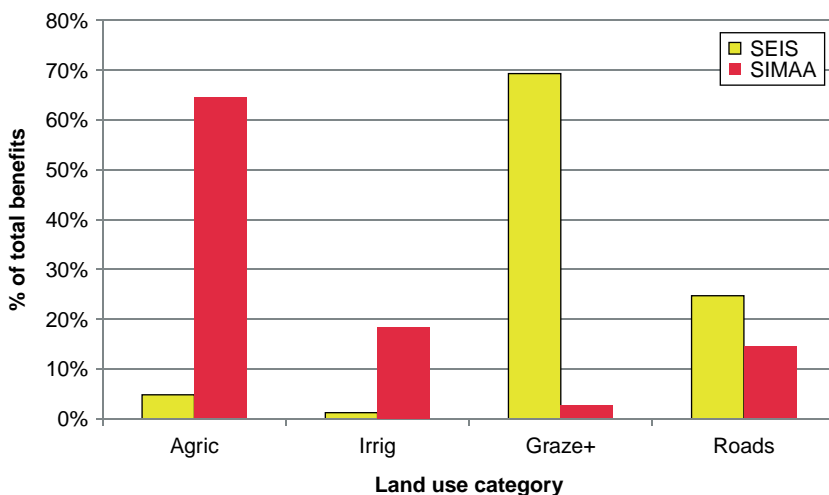
But is such an approach truly effective in contributing to secure livelihoods? Its allure rests on multiple assumptions; the land will be used, in the manner expected, by the intended beneficiaries. But these assumptions immediately raise some difficult questions, including:

- Will the beneficiary households possess the complementary inputs needed (seeds, draught animals, etc.) and have access to markets to profit from the land in the expected way?

- Do rural communities depend on access to a variety of land types to sustain both crop agriculture and livestock?

- How accurate an assessment of likely benefits will mine action NGOs be able to make?

As well, would it not be beneficial to grant demining managers flexibility in adjusting priorities to better coordinate with local development projects? The potential benefits of, say, demining in support of the rehabilitation of irrigation systems may outweigh the risk of pushing responsibility, hence discretion and the opportunity for corruption, to lower levels (Organisation for Economic Cooperation and Development [OECD] 2012).

Finally, there is the obviously point that clearing landmines is inherently about land. In rural Afghanistan, this raises the question of land disputes (Rashid et al. 2010; Szilard & Yasin 2012), which are widespread and a common source of conflicts (Wily 2004). For humanitarian actors, avoiding actions that might fuel conflict has long been enshrined in the 'do no harm' principle (Anderson 1999); a lesson now endorsed by the principal aid agencies (OECD 2010).

Clearly, there is need in Afghanistan and mine action more generally for additional evaluation tools to complement traditional project evaluations, which have been able to confirm outputs and efficiency but, due to cost and security constraints, plus the inability of donors to agree an evaluation agenda, have proved inadequate in delivering robust conclusions on livelihood outcomes, development effectiveness, and sustainability. New evaluation tools have to be appropriate for both mine action, which has its idiosyncrasies, and the challenges posed by countries such as Afghanistan, including insecurity, distrust of outsiders, and difficulty in obtaining the views of women.

> **New evaluation tools have to be appropriate for both mine action, which has its idiosyncrasies, and the challenges posed by countries such as Afghanistan, including insecurity, distrust of outsiders, and difficulty in obtaining the views of women.**

Special Features of Mine Action Affecting Evaluation Design and Implementation

Mine action has a number of features that, collectively, define it as a sector and pose additional challenges for evaluators. First, mine action comprises a variety of 'pillars': humanitarian demining, mine risk education, stockpile destruction, victim assistance and advocacy. This amalgam means mine action represents a microcosm of the broader international response to conflict-affected states. The practitioner community is dominated by military engineers, whose outlooks reflect both their professional training and the command-and-control approaches of their employers, but also includes: aid workers imbued with humanitarian principles; health care personnel working to save

lives and limbs; and information professionals wrestling with how to turn sparse and suspicious data into usable information. These professions have different ways of defining problems, programmes, and priorities; in this sense, the mine action community is a 'forced marriage' (Paterson 2004).

In addition to practitioners, there are disarmament and human rights campaigners, plus representatives from official donor agencies and the UN trying to square the circle among humanitarian imperatives, national interests, and political realities. This last feature is particularly vexing in mine action, which is both blessed and burdened by international treaties. The 1997 Ottawa Treaty and the 2008 Oslo Treaty ban the use of anti-personnel landmines and cluster munitions, respectively.[6] They also establish obligations on states parties, including the requirements to locate and clear all 'known mined areas' and 'cluster munition contaminated areas'.

Understandably, many mine action donors allocate funds not only on the basis of humanitarian needs and development opportunities, but also on each country's perceived performance in terms of the treaties. Efforts to comply with treaty obligations can distort performance in terms of reducing casualties and poverty because funds are targeted to, say, anti-personnel landmines (covered by a treaty) when UXO and anti-tank mines may pose a bigger problem.[7]

Mine action also cuts across all stages of international support for transitions from conflict to development: humanitarian, peacekeeping/stabilisation, reconstruction, and development. Mine action actors must adapt their objectives and priorities as a country transitions from conflict (often, two steps forward, then one back), and they may deliver simultaneous support to humanitarian, peacekeeping, reconstruction, and development activities. A stylised picture of this 'contiguum' (Verband Entwicklungspolitik deutscher Nichtregierungsorganisationen [VENRO] 2006, 3–4) is provided in the mine action 'programme life cycle', Figure 2 (Geneva International Centre for Humanitarian Demining [GICHD] 2004, 17–21). As a result, evaluations of mine action often

Figure 2: The Mine Action Programme Life Cycle

77
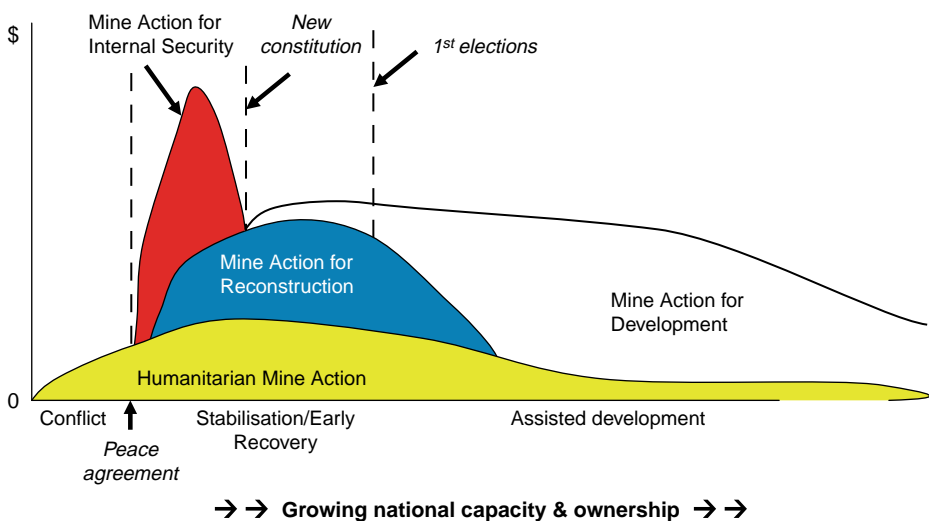
incorporate additional criteria, including safety (for deminers, aid workers, and end-users of cleared land), compliance with treaty obligations, and coverage, coherence, and coordination (OECD/Development Assistance Committee [DAC] 1999).

Insecurity is a particular concern for evaluators in Afghanistan. Security has deteriorated steadily since 2003. More than in most conflict-affected states, humanitarian space is under threat because major donors are also belligerents, because Western militaries use aid to an unprecedented extent as part of their counterinsurgency efforts, and because the UN is seen as supporting one side in the conflict (Benelli et al. 2012, 25).

Mine action operators have not curtailed their activities as much as most aid organisations. MAPA organisations operated successfully under the Taliban regime and, in recent years, have adopted community-based strategies (e.g. engaging people from local communities as deminers) that allow operations in high-risk districts. Although over 14,000 people work in mine action, they were involved in only 59 of 19,000 incidents recorded by UN in 2010–2011 (MACCA 2011, 45–47). Still, the association of mine action with the UN, targeted killings of Taliban commanders and their replacement by younger fighters who are unaware of their leadership's support for demining, and the fact that many warlords are engaged in criminality, such as kidnapping, rather than insurgency have led to an increase in abductions and attacks against mine action personnel.

Another feature of Afghanistan that evaluators need to consider is the country's geographic, climatic, ethnic, socio-economic, and cultural diversity, as well as the differential effects of the ongoing conflict. What works in one part of the country or agro-climatic zone may not hold for others. Recently, there has been a large evaluation of the huge National Solidarity Program that covered the bulk of the country (Beath et al. 2012), but time and cost constraints mean most programmes must have more modest evaluation ambitions.

The L&L Evaluations

Background

In 2008, the MACCA and the GICHD collaborated on the formulation of a new strategy for the Mine Action Program (MACCA 2008). Subsequently, MACCA and GICHD initiated discussions on an M&E system to support implementation of the strategy. They concluded that the MAPA has a well-developed monitoring system that gives a reasonable picture of what is happening (hazards and blockages of livelihoods assets; physical progress; number of beneficiaries; etc.), but provides limited insight into why things happen the way they do or how things might be done better. Therefore, the focus should be on evaluation tools to address the 'why and how questions'. A suite of tools would be required to meet the accountability requirements to donors and the government, and to address the learning agenda.

On its own, MACCA developed accountability tools including:

- Balanced Scorecard to assess operators on operational efficiency, quality management, safety and reporting;

- Project Monitoring Tool to enhance reporting to donors on progress and achievements from specific projects;

- short pre-clearance and post-clearance community surveys, providing 'post-clearance impact assessments';

- support to the government's Department for Mine Clearance (DMC) for an annual 'Post-Clearance Audit' of tasks, from which DMC would prepare reports to the government.

The GICHD contributed with a series of training workshops to support design, monitoring, and evaluation capacities within the MAPA, and with the design and delivery of 'landmine and livelihoods surveys': mixed-method community surveys employing the SL approach (Chambers 1987; Department for International Development [DFID] 1999). This is a household-level approach that places people and their priorities at the centre of development. It views people as operating in a context of vulnerability (conflict, climatic shocks, etc.), but with access to human, social, natural, financial, and physical assets. The SL approach fits naturally with mine action: mine/UXO contamination increases vulnerability because it prevents access to essential livelihoods assets.

The SL approach fits naturally with mine action: mine/UXO contamination increases vulnerability because it prevents access to essential livelihoods assets.

As illustrated in Figure 3, the vulnerability factors plus the political and institutional environment influence household strategies in pursuit of livelihood objectives. Sustainable Livelihoods Analysis is a method within the SL approach for developing a holistic understanding of the resources available to individuals, households, and communities, and the constraints and opportunities of using these resources for development; SL Analysis helps answer the 'why and how questions'.

Sustainable Livelihoods Analysis uses a mix of participatory appraisal tools to understand the local situation in a short period of study. This means working intensively with the local community. On arrival it is critical to engage with the community leaders to gain legitimacy, protection, and support in the task. It is also important to clearly explain the purpose of the study and how the findings will be used to benefit the country and the affected communities — without raising expectations. The survey team must be

Figure 3: Sustainable Livelihoods Framework

79
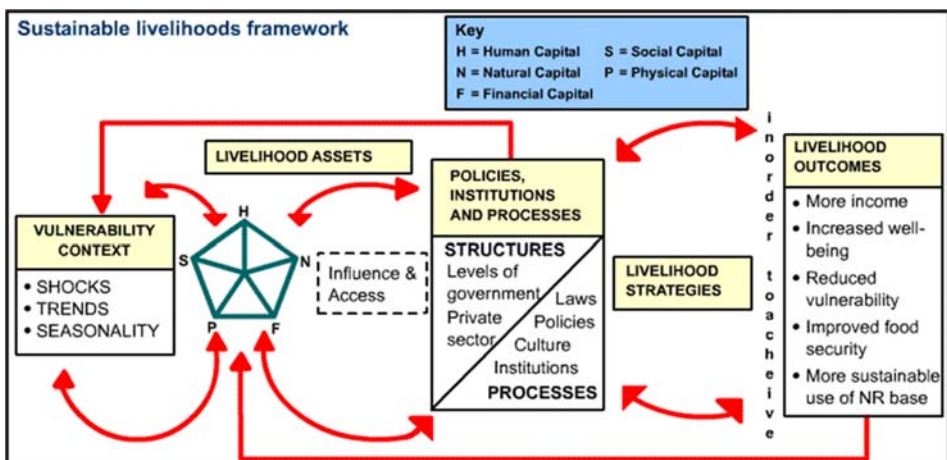
scrupulous in observing local customs and showing respect to all members of the community, by not pushing sensitive topics, and by providing a voice to the normally unheard within these communities, including specific constituencies such as women, girls, and boys.

An important challenge is to distinguish among the immediate *outputs* of mine action (e.g. cleared land and other assets; greater awareness of mine risks; survivor support), *outcomes* (e.g. utilisation of assets; higher productivity; changed behaviour); and, ultimately, *impacts* in terms of fewer casualties, sustained growth, and enhanced well-being.

GICHD had worked with the Natural Resources Institute to pilot the use of L&L evaluations in Yemen (Pound et al. 2006). To start in Afghanistan, MACCA and GICHD agreed an initial survey with a number of objectives, including:

1. Learning — to better understand the development outcomes and impacts stemming from demining and how to enhance these through:

   - adapting the priority-setting process;
   - enhancing linkages with rural development organisations.

2. Accountability — better reporting to government and donors on the contribution made by MAPA to Afghanistan's development.

3. Capacity development — ensuring the MAPA, in partnership with Afghan livelihoods experts, can conduct and analyse such surveys on a periodic basis.

The capacity development objective was central. The diversity among Afghanistan's rural communities means that no single evaluation could provide a representative picture of the mine/UXO problem and the benefits of mine action, so individual surveys should focus on specific questions on the research and evaluation agenda. MAPA organisations had a long history of conducting surveys to obtain data on both contamination and its impacts, so the assumption was that the capacity to analyse such data was the binding constraint. GICHD believed MACCA did not require social scientists on staff for this purpose, and that partnering with an Afghan policy research institute would be more cost effective and sustainable. Accordingly, GICHD signed an agreement with the government's Afghanistan Institute for Rural Development (AIRD) to contribute social scientists to support the training, survey, and data analysis. The hope was that AIRD could field a male–female team, but it was unable to recruit and retain women in such positions so the GICHD engaged a female international specialist to work with the survey leader. Four of the MAPA implementing partners each contributed a two-person (male–female) survey team.

2010 survey

Following a preparatory mission in February to agree the sampling strategy and review security requirements, the international experts delivered a five-day workshop covering theory with practical exposure to the approach, method, and tools to be used. Sustainable Livelihoods surveys draw on the Participatory Rural Appraisal toolkit (Food and Agriculture Organisation [FAO] 1999), including: secondary data analysis; timelines; maps drawn up with villagers to show the relationship between village features and

mined areas; community profiles listing community assets and external relationships; focus group discussions with community leaders, farmers, women, and children; case studies of mine/UXO survivors; gender analysis (relative to mine action); farming system diagrams; participant observation; photographic record; and qualitative vulnerability assessment of each community.

A short questionnaire was used to capture economic benefits of mine clearance (e.g. increase in land value; building of houses, schools, and businesses), but the emphasis of the L&L survey was on the qualitative outcomes of mine action. These data complement other surveys conducted by MAPA organisations which quantify outputs (e.g. area cleared).

The survey was conducted in 25 villages across four provinces to give a contrasting sample of: cleared versus partially cleared communities; different agro-ecological zones; a mix of contamination (UXO and/or mines); and peri-urban versus rural. The survey was conducted by four five-person teams (social scientist, MACCA surveyor, male–female survey team, driver) and supported by MACCA, DMC, and the male and female international specialists. As required in Afghanistan, female surveyors were accompanied by chaperones.

Normal protocol for participatory surveys would see prior consultation with local governments and advance notification to community leaders as a courtesy and to ensure a representative sample of community members would be available. Because of the dangers of abduction or other incidents, these protocols had to be modified. In addition, the length of time spent in any one community was capped at two days to reduce the risk of abduction. The teams also had to conform to a curfew, precluding travel after dark. As a result, the teams had to operate with whomever was available on the day, rather than being able to pre-arrange focus groups with participants selected according to good sampling procedures. By UN rules, international evaluators had to travel in armoured vehicles when unmarked vehicles would have been received with less suspicion in the communities. Despite the security restrictions, the survey teams were able to keep to their planned itinerary.

Surveyors worked in Dari, requiring translation before the analysis and reporting in English. The English-language report incorporated an 'extended Executive Summary' in Dari, which was distributed among stakeholders (MAPA organisations, donors and UN agencies, government officials, and some development NGOs). A stakeholders' workshop was held to review and validate the findings, conclusions, and recommendations, after which the report was finalised. The report documented that demining freed a wide variety of livelihoods assets, creating opportunities that enabled the following developmental outcomes:

Physical security: MACCA data showed 363 total casualties before demining in the communities surveyed, while no community reported post-clearance casualties from mines/UXO. This commendable record provided great relief, particularly for women ('The benefit of demining is that we feel safe: if our children go out of the house or our husbands go to work we feel relaxed because they are safe' — woman, Ala Chapan community).

> **MACCA data showed 363 total casualties before demining in the communities surveyed, while no community reported post-clearance casualties from mines/UXO.**

Most community members reported confidence that areas were safe after clearance for agriculture, grazing, recreation, passage, and construction. Areas with economic or cultural value were utilised quickly after clearance. However, the survey also highlighted that there was no explicit focus on community liaison by deminers, and no systematic approach to ensure that women were included in these discussions (e.g. so they know a hazard has been cleared).

Women carry a greater psychological burden, in spite of much lower exposure to risk, because of their seclusion and dependence on second-hand information.

An interesting result, depicted in Figure 4, was that women perceived the danger from mines/UXO to be much greater than that reported by men or according to MACCA's casualty data. Women carry a greater psychological burden, in spite of much lower exposure to risk, because of their seclusion and dependence on second-hand information.

All villages surveyed had received at least some mine risk education (MRE), adult males and children reporting that they had received MRE more often than did adult women. However, coverage appears to be far from universal. Not all children attend school where MRE is delivered, and many women have restricted mobility and cannot attend meetings.

Social outcomes: Communities reported greater use of recreational areas, plus (re-) construction of mosques, schools, community centres, and other social amenities.

Humanitarian outcomes: The support provided to mine/UXO victims was extremely gender-biased, men being much more likely to receive artificial limbs, government disability pensions, or loans to start businesses. Only one example of a woman receiving such victim assistance was identified. Both male and female survivors received free medical treatment in most cases, but treatment depended on getting to a hospital, which is difficult for people in remote villages.

Economic outcomes: The absence of post-clearance casualties leads to reduced medical costs and increased productive labour. The survey also recorded significant economic

Figure 4: Perceptions of Danger from Mines/UXO
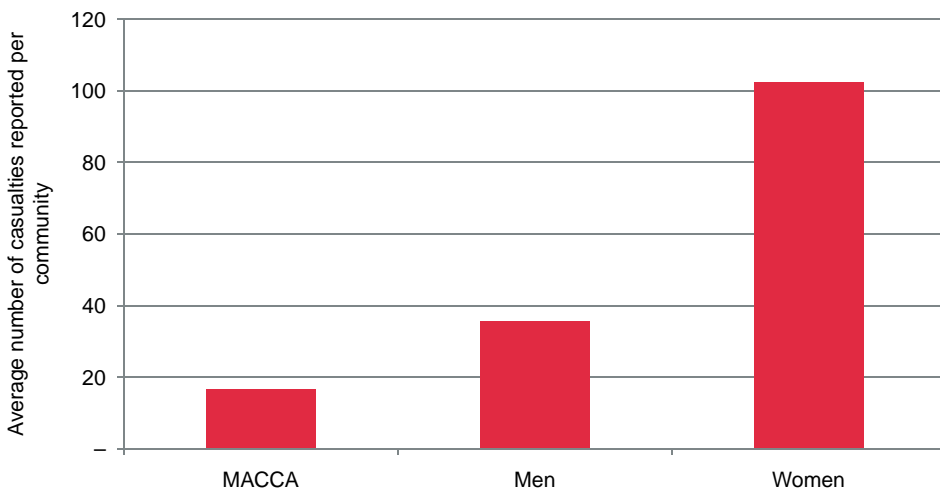
activities on cleared land, both for the community and for Kuchi pastoralists. Cleared land is mostly returned to its rightful owners (government, private, or communal ownership) and quickly used for productive purposes, demonstrating the essential nature of these assets. In a minority of cases, villagers were unhappy about the way in which the land had been allocated (e.g. land grabbing by local politicians). In some cases, minefield clearance is a sound economic investment, even before considering humanitarian benefits or treaty obligations.

Strategic outcomes: Demining allowed projects of national importance to proceed and the return of migrants, refugees, and internally displaced people.

Additional development opportunities: During separate focus groups, men and women were asked about the additional developments that would most benefit their community. While men emphasised productive opportunities made possible by clearance, women consistently highlighted safety and recreational benefits that gave them peace of mind and a better life for their children.

Many communities stated that lack of assistance from government agencies or NGOs was holding up the use of agricultural assets because, for example, funds and expertise were required to rehabilitate irrigation works.

The prioritisation process: Prioritisation of demining in Afghanistan is based on specified indicators of the harmful impacts of contamination (recent accidents, blocked access to crop land, water sources, infrastructure, etc.) and opportunities (e.g. resettlement of displaced persons, community development projects). Each indicator is weighted to reflect its perceived importance. After applying these weighted indicators, each hazard is classified as high, medium, or low impact. Hazards with recorded victims and that block resettlement are automatically classified as high impact (GIRA 2012, 185–186).

The survey found that villagers are largely satisfied with the prioritisation of hazards within their communities and that, in most cases, the priority of villages in terms of demining was deemed appropriate. In many cases, community men were pleased to have been involved in discussions about the order in which hazards were cleared. Women were much less likely to have been involved.

Survey quality and replicability: This initial survey was a pilot to test tools and develop the capacities of local organisations and social scientists. Via participatory capacity assessment at the end of the survey, the surveyors reported confidence in their ability to conduct similar surveys with the support of AIRD. Key lessons included:

- Including women surveyors considerably enhanced the breadth of information obtained.
- The use of a range of participatory tools allowed good triangulation.
- During the survey, opportunities organised for surveyors to interact within and across teams to compare notes and form a full picture of the findings proved very successful.
- The method and tools were culturally appropriate, although additional tools, such as daily and seasonal calendars for women, would provide useful information on their exposure to risk.

- Due to the skill of MACCA and DMC in monitoring the situation, and to the esteemed status of deminers, there were no security problems and the teams were welcomed into, and courteously attended by, all communities.

As expected, shortcomings were also identified:

- The patchy quality of the data collected suggested a need for further training and experience, particularly in probing (asking follow-up questions to obtain understanding in depth and breadth). While some useful financial data were collected, they often were insufficient for economic analysis.

- The translation of village datasets from Dari took a long time and significant detail was lost in the process.

- The link with AIRD for social science expertise was an excellent initiative and, together with the logistical and technical expertise of MAPA, provides the basis for an in-country capability for future evaluations. However, the value of the link was later reduced when both individuals left AIRD for alternative employment.

- The failure of the UK to issue visas meant MACCA and AIRD personnel could not participate directly in the data analysis and drafting the report, and an important opportunity for capacity development was lost.

- The security situation imposed restrictions on the survey. UN regulations required armoured vehicles for international personnel and locations were chosen partly on the basis of their relative security. Also, time spent in communities was restricted and teams could not spend the night in villages to develop a closer relationship with community members. By developing the capacity of local personnel, some of the security restrictions could be lifted, providing the opportunity for a wider sample of, and greater engagement with, communities.

2011 survey

Given the successes and shortcoming of the initial survey, MACCA and GICHD agreed on a second, smaller survey focusing on capacity development and verifying findings in a different province. The aim was to ensure that MAPA, in partnership with Afghan social scientists, could conduct such surveys on a periodic basis and analyse the data using the SL model. Additional training (including in probing, the use of daily and seasonal calendars, and gender analysis) was provided. New male and female social scientists were assigned who had limited experience with the SL approach. Therefore, training also had to bring these social scientists up to speed.

The methodology originally proposed the use of wealth ranking as a tool to differentiate outcomes for households of different wealth status, but this had to be dropped because of concerns it might identify individuals who could then be at risk of abduction for ransom.

Data from the second survey were more comprehensive, due partly to the stronger women's teams (which included two female Afghan social scientists in addition to an international specialist), and partly to experience from the first survey. Data analysis and report writing were done jointly by Afghan and international staff to develop capacity in this key area.

One interesting finding demonstrated the marked differences in the responses by male and female focus groups. While the men in one community did not mention land disputes, although the question was asked explicitly, the women from the same

community reported many disputes. When asked the reason, the women said that clearance resulted in many people being interested in buying the safe land. Some landowners capitalised on this demand by selling the same plot to as many as four buyers. According to the women, there were about 100 cases where one piece of land was sold to more than one person.

Capitalising on L&L Evaluations

What was learned?

The surveys broadly confirmed that demining has led to substantial humanitarian, social, and economic benefits. In many cases, however, sustained benefits from the freed assets are contingent on additional investment by households, the community collectively, or development agencies. The findings confirm that efforts to strengthen coordination between communities and both government departments and development NGOs represent a sound investment. The findings have also contributed to an understanding by MACCA that additional efforts are required to ensure the views of women are obtained, and to determine whether demining might aggravate land disputes.

> **The findings have also contributed to an understanding by MACCA that additional efforts are required to ensure the views of women are obtained, and to determine whether demining might aggravate land disputes.**

The involvement of internationals requires measures to reduce security risks, leading to limitations on travel, overnight stays in villages, and modes of transport; restrictions that impinge on the trust established with, and the information provided by, communities. However, the initial surveys helped develop the capacity of MACCA and AIRD to the point where they were able to plan, conduct, analyse, and report L&L evaluations on their own. Continued support from international organisations can ensure that capacity development continues and leads to further enhancements of L&L surveys, but direct participation by international experts in community visits is no longer required.

L&L surveys are relatively expensive when the direct participation of international specialists is required. The cost of the 2010 survey project, including training, the stakeholder workshop, and the time of GICHD and MACCA personnel, was about US $200,000, while that of the smaller 2011 survey was perhaps US$100,000. By comparison, a survey without direct involvement of international personnel would be about US $30,000.[8]

But any additional cost would be hard to justify if the L&L evaluations were not useful. Fortunately, MACCA has implemented key recommendations and developed its capacity for evaluation in support of its learning agenda.

Applying what was learned

Since the completion of the first L&L survey, MACCA has implemented key recommendations and further developed evaluation capacities in support of its learning agenda. It established an R&D unit focusing, in part, on the design, conduct, analysis, and reporting of such surveys. In September 2012, this unit conducted a third L&L evaluation in Badakhshan Province. This survey also involved female surveyors

from demining organisations, as well as male and female consultants from AIRD. Future evaluations are planned with the support from AIRD.

Other MAPA organisations are also making use of their L&L survey experience. For example, HALO Trust created two male–female teams to conduct community livelihoods surveys. Its teams are using the same approach and questionnaires used in the MACCA evaluations, in which HALO's male and female surveyors participated.

The DMC also engaged a female surveyor, for the first time, to work on its post-clearance audit for 2012. Unfortunately, she was unable to join the audit team until the end of the exercise, but the initiative bodes well for future years.

In other cases, MACCA's experience with L&L surveys has contributed, along with other factors, to the launch of new initiatives. For example, given the dramatic differences in the information obtained from women and men in communities visited during the L&L surveys, MACCA plans to conduct a gender assessment during 2013. This will be done by an international specialist to assess whether MACCA's internal practices and programme-wide policies to support gender mainstreaming are well conceived and being followed.

MACCA has also taken steps to strengthen its coordination with ministries and other development organisations to enhance mine action and development linkages. There now are regular meetings to discuss each ministry's priorities. MACCA provides technical advice and maps showing the location of any cleared areas and remaining hazards close to a project site. Since 2011, ministries and other development organisations have shared plans for 368 projects with MACCA. MACCA found that mine/UXO contamination will create implementation problems for 42 projects. Subsequently, MACCA tasked demining assets to remove hazards affecting seven projects and, for the rest, provided technical advice to the ministries to revise project plans or solve the problem using demining firms.

Finally, MACCA has strengthened procedures to avoid land disputes arising from demining. In coordination with the communities, the possibility of land disputes is checked in the initial stage of developing demining projects. There now is as well an unequivocal statement in the handover certificate issued after each hazard has been cleared that indicates the certificate is not a document that confers land rights; rather it certifies only that the area cleared is free of landmines/UXO and safe for use. As a result, those receiving handover certificates cannot use the documents in support of claims to, for example, communal land.

Conclusion: Lessons for Evaluation in Violently Divided Societies

The experience from Afghanistan suggests that many participatory rural appraisal tools are appropriate in violently divided societies. Communities everywhere are sensitive, unpredictable entities and surveyors must be culturally aware and expect site-specific peculiarities that require adaptation of survey tools. In all communities there are different perspectives. Participatory methods allow these to be identified and explored. In one sense, a violent history is simply an additional dimension of such perspectives.

However, obtaining in-depth understanding was extremely difficult given resource and security constraints. UN regulations prevented teams from spending the night in the villages to build trust and provide a safe space for informal discussion and storytelling, and teams could only spend about two hours with any one group — hardly adequate to explore deep and painful experiences. The use of local surveyors and social scientists reduces the need for such restrictions, facilities the emergence of trust, and leads to greater understanding. The greater use of local evaluators provides at least a partial solution to the evaluation challenges that arise in violently divided societies, and highlights the need for greater evaluation capacity development efforts within these countries.

> **UN regulations prevented teams from spending the night in the villages to build trust and provide a safe space for informal discussion and storytelling, and teams could only spend about two hours with any one group — hardly adequate to explore deep and painful experiences.**

In addition, those planning evaluations must be attuned to the dynamics of violence in each country, and adapted accordingly. The involvement of local experts makes it more likely that appropriate adjustments will be made. For example, the proposal by the evaluation team leader to use a wealth-ranking tool was dropped because Afghan social scientists were alert to the growing risk of abduction for ransom.

A sense of realism is also needed when conducting evaluations in violently divided societies. In the L&L evaluations, surveyors were aware that, in some communities, they obtained a sanitised version of the truth. Teams visited communities where there was smuggling or poppy production, but these were not mentioned. In others, surveyors observed symptoms of intra-community conflict and the influence of power, but these topics were not volunteered during interviews. In one community the men did not mention the serious land disputes: land disputes were only confirmed by the women.

This last example also highlights the significant and consistent differences between men's and women's responses. These gendered perspectives justify the expense and logistical challenges of having both male and female teams and underline the need to ensure that women are included in evaluation capacity development initiatives.

TED PATERSON is a development economist who, after 15 years working with NGOs and policy research institutes, wandered into mine action in 2000. Currently, Ted is Senior Advisor for Strategic Management at the Geneva International Centre for Humanitarian Demining.

BARRY POUND started his career as a tropical agronomist, later focusing on farming systems and livelihoods. He worked with the Natural Resources Institute (University of Greenwich) for 20 years. Now a freelance consultant, Barry was survey team leader for the L&L surveys in Yemen and in Afghanistan.

ABDUL QUDOS ZIAEE is an engineer who joined mine action in 1999 as a demining surveyor, then moved to MACCA. Qudos contributed to the first two L&L surveys in Afghanistan, and now is the task manager for these surveys at MACCA.

Endnotes

[1] Humanitarian demining implies survey and clearance activities that are not for military or purely commercial purposes. It is not limited to demining in support of humanitarian aims.

[2] See Pound et al. 2011, 2012.

[3] MACCA is a project of the UN Mine Action Service (UNMAS) and serves as the de facto national mine action centre on behalf of the government.

[4] Mine action is also financed by governments in mine-affected countries (US$195 million recorded by Landmine Monitor for 2011), by the UN assessed budget for peacekeeping operations (almost US$90 million), and via commercial contracts for public and private investments, which amounted to an estimated US$85 million in 2009–2010 for Afghanistan alone. Globally, non-military mine action is at least a US$1 billion per year industry.

[5] For the case of Sudan, see Bennett et al. 2010.

[6] The official titles are Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on Their Destruction (Ottawa Treaty) and the Convention on Cluster Munitions (Oslo Treaty). Although some states have not ratified these conventions, they have been successful in establishing international norms against the use of such weapons, and many non-signatories abide by those norms.

[7] 'From a humanitarian perspective, Afghanistan cannot focus only on AP removal at the expense of AT and [battle field] removal. There are AT minefields and [battlefields] with a higher priority for clearance than some AP minefields' (GIRA 2012, 9). UXO — not landmines — accounts for about two-thirds of civilian casualties in Afghanistan.

[8] Another way of looking at the value-for-money of L&L evaluations is a 'reasonableness comparison': the cost of evaluation versus the cost of mine action. Using clearance data from MACCA, about US$30 million had been spent on demining in the 25 communities covered by the 2010 survey. The pilot L&L evaluation cost only 0.66% of the demining expenditures and significantly added to the understanding of the outcomes achieved.

References

Anderson, M. 1999, *Do No Harm: How Aid Can Support Peace — or War*, London: Lynne Rienner.

Andersson, N., **Whitaker, C.** & **Swarminathan, A.** 1998, *Afghanistan: the 1997 National Mine Awareness Evaluation (MAE)*, Ottawa: CIET for UNOCHA.

Beath, A., **Christia, F.** & **Enikolopov, R.** 2012, *Randomized Impact Evaluation of Afghanistan's National Solidarity Programme — Final Report*, Kabul: Ministry of Rural Rehabilitation and Development and World Bank, http://www.nsp-ie.org/endline/reportsfinal.html, accessed 29 August 2012.

Benelli, P., **Donini, A.** & **Niland, N.** 2012, *Afghanistan: Humanitarianism in Uncertain Times*, Boston: Feinstein International Centre.

Bennett, J., **Pantuliano, S.**, **Fenton, W.**, **Vaux, A.**, **Barnett, C.** & **Brusset, E.** 2010, *Aiding the Peace: A Multi-donor Evaluation of Support to Conflict Prevention and Peacebuilding Activities in Southern Sudan 2005–2010*, ITADLtd, UK: http://www.oecd.org/countries/southsudan/46895095.pdf, accessed July 2013.

Byrd, W.A. & **Gildestad, B.** 2001, *The Socio-economic Impact of Mine Action in Afghanistan: A Cost–Benefit Analysis*, Washington: World Bank, http://www.gichd.org/lima/reports-publications/detail/publications/the-socio-economic-impact-of-mine-action-in-afghanistan-a-cost-benefit-analysis/

Chambers, R. 1987, 'Sustainable Livelihoods, Environment and Development: Putting Poor Rural People First', IDS Discussion Paper 240, Institute for Development Studies, Brighton.

Convention on Cluster Munitions (Oslo Treaty) 2008, http://www.clusterconvention.org/

Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on Their Destruction (Ottawa Treaty) 1997, http://www.apminebanconvention.org/

Christoplos, I. 2004, *Out of Step? Agricultural Policy and Afghan Livelihoods*, Kabul: Afghanistan Research and Evaluation Unit.

DFID 1999, 'Sustainable Livelihoods Guidance Sheets', London: www.eldis.org/go/topics/dossiers/livelihoods-connect/what-are-livelihoods-approaches/training-and-learning-materials

FAO 1999, PRA Tool Box, Module 6 of 'Conducting a PRA Training and Modifying PRA Tools to Your Needs', Rome: http://www.fao.org/docrep/003/X5996E/X5996E00.HTM

GICHD 2004, *A Guide to Socio-Economic Approaches to Mine Action Planning and Management*, Geneva: http://www.gichd.org/publications/year/2004/a-guide-to-socio-economic-approaches-to-mine-action-planning-and-management-en

Gildestad, B. 2001, *Study of Socio-economic Impacts of Mine Action in Afghanistan*, Islamabad: World Bank and UNDP.

GIRA 2012, 'Request for an Extension of the Deadline for Completing the Destruction of Anti-personnel Mines in Mined Areas in Accordance with Article 5 of the Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on Their Destruction', revised request submitted 31 August, Kabul: http://www.apminebanconvention.org/states-parties-to-the-convention/afghanistan/, accessed 26 December 2012.

ICBL 2012a, *Landmine Monitor 2012*, Ottawa: Mines Action Canada, http://www.the-monitor.org/index.php/publications/display?url=lm/2012/

ICBL 2012b, *Landmine and Cluster Munition Monitor: Country Profile for Afghanistan 2012*, Ottawa: Mines Action Canada, http://www.the-monitor.org/index.php/cp/display/region_profiles/profile/413, accessed 30 December 2012.

MACCA 2008, 'Mine Action Programme for Afghanistan: Mine Action Strategic Guideline 2008–2013', Kabul: www.macca.org.af/en/mapa_plans.html, accessed 5 May 2010.

MACCA 2011, 'MAPA Annual Report 1389', Kabul: http://www.macca.org.af/en/MAPA_Reports.html, accessed 12 June 2012.

Maletta, H. 2004, *The Grain and the Chaff: Crop Residues and the Cost of Production of Wheat in Afghanistan in a Farming-System Perspective*, Buenos Aires: Universidad Nacional Tres de Febrero, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=571343

MCPA 2001, *Socio-economic Impact Study of Landmines and Mine Action Operations in Afghanistan*, Kabul.

OECD 2010, *Do No Harm: International Support for Statebuilding*, Paris.

OECD 2012, *Managing Risks in Fragile and Transitional Contexts: The Price of Success?*, Paris.

OECD/DAC 1999, *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies*, Paris.

Paterson, T. 2004, 'Time to Go MAD: Mine Action and Development' in Maslen, S. ed, *Mine Action after Diana: Progress in the Struggle Against Landmines*, London: Pluto Press.

Paterson, T. 2005, 'Is Mine Action Making a Difference ... or Avoiding the Question?' in GICHD, *Mine Action: Lessons and Challenges*, Geneva: http://www.gichd.org/publications/year/2005/mine-action-lessons-and-challenges-en

Pound, B., **Martin, A.**, **Qadr, A.** & **Mukred, A.W.** 2006, *Departure of the Devil: Landmines and Livelihoods in Yemen: Volume 1*, Geneva: GICHD, www.gichd.org/strategic-management/mine-action-security-and-development/socio-economic-survey/landmines-and-livelihoods

Pound, B., **Massleberg, Å.**, **Ziaee, A.Q.**, **Hashimi, S.**, **Qayyumi, S.** & **Paterson, T.** 2012, *Livelihoods Analysis of Landmine/ERW Affected Communities in Herat Province, Afghanistan*, Geneva: GICHD,

www.gichd.org/strategic-management/mine-action-security-and-development/socio-economic-survey/landmines-and-livelihoods

Pound, B., **Wood, A.**, **Farahi, S.Z.**, **Popal, R.**, **Paterson, T.**, **Ziaee, A.Q.** & **Hashimi, S.** 2011, *Livelihoods Analysis of Landmine Affected Communities in Afghanistan*, Geneva: GICHD, www.gichd.org/strategic-management/mine-action-security-and-development/socio-economic-survey/landmines-and-livelihoods

Rashid, M.S., **Jan, M.** & **Wakil, M.** 2010, *Landmines and Land Rights in Afghanistan*, Geneva: GICHD, http://www.gichd.org/strategic-management/mine-action-security-and-development/landmines-and-land-rights/research-and-publications/

Szilard, F. & **Yasin, S.** 2012, *Land Rights and Mine Action in Afghanistan: Analysis and Programming Entry-points for Mine Action Organisations*, Geneva: UN-HABITAT and GICHD.

Van Ree, B., **Khan, M.**, **Paterson, T.** & **Tariq, Q.** 2001, 'Report on the Review of UNOCHA's Mine Action Programme for Afghanistan', Islamabad: http://www.gichd.org/lima/reports-publications/detail/publications/report-on-the-review-of-unochas-mine-action-programme-for-afghanistan/

VENRO 2006, 'Linking Relief, Rehabilitation and Development: Approaches and Financing Instruments to Improve the Transition between Relief, Rehabilitation and Development Co-operation', VENRO Working Paper No. 17, Berlin: http://www.venro.org/fileadmin/Publikationen/Einzelveroeffentlichungen/Humanitaere_Hilfe/arbeitspapier_17_en.pdf

Wily, L.A. 2004, *Looking for Peace on the Pastures: Rural Land Relations in Afghanistan*, Kabul: African Research and Evaluation Unit, http://www.areu.org.af/EditionDetails.aspx?EditionId=600&ContentId=7&ParentId=7

90

Thinking Creatively About Methodological Issues in Conflict-Affected Societies: A Primer from the Field

Sylvester Bongani Maphosa
Published online: 19 Sep 2013.

PLEASE SCROLL DOWN FOR ARTICLE

THINKING CREATIVELY ABOUT METHODOLOGICAL ISSUES IN CONFLICT-AFFECTED SOCIETIES: A PRIMER FROM THE FIELD

SYLVESTER BONGANI MAPHOSA

Abstract

Conducting interviews and collecting data with which to evaluate peacebuilding programming may be problematic in many conflict-affected societies. However, many researchers do not document their dilemmas and chosen courses of action in data collection and analysis. This article explains the design, data collection, and analysis of a study that applied exploratory interpretive mixed methods to examine the impacts of one community-based peacebuilding intervention at the individual and village level in Burundi. It presents the procedures from field preparation to data analysis and candidly discusses the influence of methodological considerations in a conflict-affected setting on the unfolding research activity. While the approach may seem basic, this article seeks to shed light on the methodological choices faced by peacebuilding evaluators, thereby serving an important pedagogical function, and is intended equally for the benefit of experienced as well as novice researchers.

Keywords: exploratory interpretive mixed methods, research design, case description, focus groups, methods and procedures, research activity, data collection, group dynamics, peacebuilding impact, data analysis

Introduction

Many researchers do not document the dilemmas they face and how they arrived at their chosen courses of action during the data collection and analysis phases of peacebuilding evaluations. Moreover, conducting interviews and collecting data with which to evaluate peacebuilding programming are often problematic in conflict-affected societies. Several intersecting dynamics make peacebuilding evaluation a thorny undertaking, including: (i) the conflict context; (ii) the evolution of peacebuilding as a normative and programmatic agenda; and (iii) the limitations of evaluation methodologies in establishing correlative or causal links between interventions and outcomes. For example, Menkhaus (2004) demonstrates that many problems reflect generic, endemic social science problems of measurement and attribution of causality, and others more specific to the peacebuilding project impact. Even if good indicators can be found to measure peacebuilding impact, the specific data needed to measure

> **Many researchers do not document the dilemmas they face and how they arrived at their chosen courses of action during the data collection and analysis phases of peacebuilding evaluations.**

> **The process of gathering data requires relatively high degree of resource capability and can result in evaluation outcomes that are patently obvious and do not capture the nuances of conflict transformation and community regeneration needed to understand what does and does not work — and why.**

those indicators can be very difficult to secure in a post-conflict setting (Church & Shouldice 2002; 2003). The process of gathering data requires a relatively high degree of resource capability and can result in evaluation outcomes that are patently obvious and do not capture the nuances of conflict transformation and community regeneration needed to understand what does and does not work — and why.

Recent scholarship has attempted to address these field challenges in many ways. Some have postulated new approaches to conceptualising the meaning of success in community regeneration and peacebuilding efforts (D'Estree et al. 2001; Kriesberg 2007; Mitchell 1993; Ross 2004; Rouhana 2000). Others demonstrate the conditions that enable 'good enough' engaging dividends (Douma & Klem 2004; Lieberfeld 2002). Still others have suggested useful frameworks that could be applied to unlock the transformation potential undertaking of peacebuilding and conflict prevention (Anderson & Olson 2003; Church & Shouldice 2002; 2003; Paffenholz & Reychler 2007). Yet despite this impressive body of scholarship, many practitioners and academics remain hesitant about the overall utility of peacebuilding evaluation. Anderson and Olson (2003) have argued that the complexity of peacebuilding work, especially in fragile theatres, makes data collection and impact evaluation extremely difficult to conduct.

This article presents the considerations and procedures from field preparation to data analysis for one peacebuilding evaluation in rural Burundi. In doing so, it contributes to the evaluation literature, expanding the documentation on procedures and experiences of data collection and analysis in conflict-affected situations. By deeply engaging with these fundamentals, the article seeks to shed light on the methodological choices faced by peacebuilding evaluators, thereby serving an important pedagogical function, and is intended equally for the benefit of experienced as well as novice researchers.

The Research Design and Case Description

The peacebuilding evaluation that this study chronicles sought to ascertain the impact of peacebuilding work performed by Centre Ubuntu (CU) in three villages in rural Burundi. The epistemological orientation of the study was essentially exploratory, and so employed a case study interpretive mixed methods design, which is thought to allow for a 'breadth and depth of understanding and corroborating' (Johnson et al. 2007, 123) the impact of CU and their peacebuilding activities (independent variables) on peace-building outcomes at the individual and village levels (dependent variables). Applied exploratory interpretive mixed methods are described as a design that employs both quantitative and qualitative methods and methodologies in a single study (Johnson et al. 2007; Johnson & Onwuegbuzie 2004). It permits the researcher not 'to be constrained to either one of the traditional, though largely arbitrary, paradigms when [I] can have the best from both' (Johnson et al. 2007, 116).

The studied evaluation had a sequential status design, meaning that it transitioned from a qualitative-dominant to a quantitative-dominant design after primary data collection. Procedures were geared towards the collection of qualitative data, which were later quantitated for statistical purposes. In effect, one type of data (narrative data) provided a

basis for another (numerical codes for descriptive statistics), informing the process of variable creation and affording explanatory insights. Data were derived from both primary (focus group discussions, key informant in-depth interviews, and observation) and secondary (library materials) sources.

This mixed method interpretive approach was chosen for a number of reasons, including to:

- encourage equal participation of all important stakeholders in a shared learning process;

- thereby contribute to the promotion of the local peace and reconciliation process (Lederach 1997, 138–140);

- assess the effect of peacebuilding efforts on individuals and community groups;

- glean useful information on results at process and output levels; and

- provide an element of triangulation, given 'the very practical nature of evaluation research and the need for multiple sources of evidence when judging social programmes' (Johnson et al. 2007, 116).

The Research Activity

The following section provides an analysis of how the conflict context — including country risks (geographical, political, security), partner risks (capacity of assistants), and operational risks (roads, transport, time, financial resources) — introduced methodological issues during the research activity. The latter included challenges with regard to moderator training, sampling, recruitment of participants, focus group conduct and composition, payment, time management, and data recording and analysis.

Quality fieldwork depends on careful planning and implementation. When I travelled to the field, the research instruments were designed and ready. I had sketched the analysis plans and identified pre-set themes for data analysis and interpretation. My most immediate goal was to cultivate trusting relationships with my two assistants and the communities where data were to be collected. I carefully considered the manner in which I

Quality fieldwork depends on careful planning and implementation.

was going to introduce myself and the study to ensure: (i) research assistants understood the purpose of the study, and (ii) successful community entry and participant recruitment.

A first priority was for the team members — principal researcher and the two assistants — to get to know each other and spell out what it was that we were to do. We took turns to describe our backgrounds and general interests. The two assistants — one female (Madeleine Bigirimana) and one male (Placide Ntahonshikiye) — were employed at CU as researcher and communications officer, respectively. It was important that I recruited a female and male team to mitigate cultural and gender barriers. I then explained their anticipated roles in the project as both research assistants and moderators. Articulating their role and function, I highlighted my expectations of them, from data collection to analysis and data interpretation.

I had them undergo a three-day orientation which I designed and facilitated. The orientation focused on learning outcomes, namely (i) to provide assistants with basic

techniques for focus group facilitation, moderation, recording group dynamics, and observation, and (ii) to internalise the skills associated with conducting focus group discussions through active role play. Table 1 lists the topics covered on each of the three days. Obtaining an accurate translation from English to Kirundi (and vice versa) proved a challenge. The translation exercise had to be done carefully because 'words carry nuances and multiple meanings that are difficult to replicate from one language to another' (Maynard-Tucker 2000, 402). We discussed the choice of words until we arrived at consensus. Importantly, the exercise was useful not only in translating the instrument into the local language, but also in deepening understanding and familiarizing the assistants with the instruments. Furthermore, by Day 2 we were slowly bonding as a team. Working in an environment like CU naturally generated keen interest from assistants to learn about mixed methods research and focus group methodology, as well as how we were going to collect data in the field. This helped to break down the barrier of power relations between myself and them. A team mindset was beginning to emerge.

Sampling and recruitment

Sampling involved selecting focus field sites and group participants. Although effects of the war are widespread in Burundi and CU has interventions in all 17 provinces of the country, three village sites were selected and represented a microcosm of the complex issues confronting Burundi: Ruziba in Kanyosha (Bujumbura province), Buhoro in Itaba (Gitega province), and Ruhororo in Ruhororo (Ngozi province). The selection criteria were strategic and included three principal issues: (i) security of the research team; (ii) accessibility in terms of time, money, and terrain; and (iii) the diversity of CU programming.

A central concern related to the representativeness of potential participants to allow the evaluation to draw generalizable conclusions. Since I sought to illuminate the processes involved in peacebuilding practice, it made sense to access participants who had direct experience of the violent past and CU programme.

Table 1: Topics Covered during the Three-Day Orientation

Day	Topics covered
1	The goals of the research How 'lessons learned' will be shared with CU and relevant stakeholders The research design Anticipated methods of data collection Development of the instruments Overview of techniques for conducting individual interviews, group interviews, and focus group discussions Translation of the English instruments into Kirundi
2	Translation of the English instruments into Kirundi
3	Conduct of focus group discussions, emphasising: Introduction of facilitators and research purpose The importance of maintaining confidentiality Obtaining informed consent (for notes, and voice and video recording) Facilitating the discussion Recording data Participant withdrawal Thanking participants for their time Role play of scenarios for focus groups' discussion

Usually, random sampling in focus group studies at best produces a broadly representative sample but is unlikely to help in recruiting the diverse range of participants required for qualitative research (Bloor et al. 2001). On this basis, it was prudent to use purposive sampling, 'which [allows] units [of] a population to be studied [to be] selected on the basis of the researcher's own understanding and knowledge of the population, its elements, and the nature of the research aims' (Babbie & Mouton 2007, 166). In this way, we speculated as to which groups of people were likely to have different views and/or experiences (Kuzel 1992; Mays & Pope 1995), such as former rebels, demobilised soldiers, internally displaced persons (IDPs), returnees, widows and other victims. We used this tagging to guide recruitment and group composition. The rationale for such a range of selection characteristics was to aggregate a group discussion that would articulate and generate best possible and relevant experiential viewpoints about the peacebuilding role of community-led intervention. Indeed, there was a danger of possible re-traumatization of the community and amplification of latent conflict dynamics. However, after carefully considering the benefits and harms the methods and procedures would produce with respect to advancing the common good of the whole of Burundi and consolidating moral virtues, I was convinced my decisions were ethically justified. If confronting the truth is critical for reconciliation, then pooling victims and perpetrators is a necessary part of dealing with the violent past. In that way, the methods and procedures become a practical vehicle to further the transformation of conflict dynamics and actors.

The rationale for such a range of selection characteristics was to aggregate a group discussion that would articulate and generate best possible and relevant experiential viewpoints about the peacebuilding role of community-led intervention.

Participant groups of men and women with direct experience of the violent past were constituted to approximate control and treatment groups: (i) direct targets of the CU programme, and (ii) those not directly targeted by or participating in the CU programme. I considered and intended to interview people who did not live in the same villages that CU was operational in so as to eliminate spillover effects of living in proximity with those who had gone through the programme. However, I did not wind up having the time and resources to include another group of study participants.

The assistants helped me to successfully enter villages in non-disruptive ways and in accordance with Kirundi social norms. This was a feat easily achieved given that CU had established healthy relationships with the villagers. So, through them, I was able to ensure a successful recruitment process and optimize diversity in participants.

We telephoned formal and informal gatekeepers in each village and explained the nature of the study, as Wilkinson (1999) suggests. Although Wilkinson did not explain why collaboration with leaders was important, in this study we solicited the permission of local leaders to conduct focus groups in their villages so as to be transparent about our intentions. While it is standard practice to ensure that researchers have permission and authority to conduct research in any particular location, consulting with gatekeepers was also an important first step in earning whole-of-community trust and ensuring honest participation. In violently divided contexts such as Burundi, obtaining entry permission is one thing; earning trust is another. One may enter but still not be trusted! Earning trust therefore was essential to ensure partnership

Earning trust therefore was essential to ensure partnership and citizen control (critical to generate valid and trustworthy data) rather than token participation.

and citizen control (critical to generate valid and trustworthy data) rather than token participation. We furnished contact point persons with a description of characteristics of the kind of participants we sought to recruit for the focus groups. Identifying the selection criteria for inclusion and exclusion was important to assist contact point persons in the recruitment. Clear understanding of the reasons why some people were included and others not included was essential to guide the compilation of a list of community members whose participation was to be solicited (Vissandjee et al. 2002).

Working with local facilitators who knew the villages and population well facilitated my acceptance by, and entry into, the communities. Furthermore, village penetration through local community gatekeepers was critical to building healthy rapport with community members and to the success of the focus group data collection. Showing respect for local gatekeepers, and having community members witness this respect, demonstrated sincere interest to engage respectfully with the community (Vissandjee et al. 2002). Throughout, I was careful not to romanticise entry and acceptance, because edicts permitting entry can sometimes be a function of power used to serve the interests of narrow groups whose communities are stifled. In that sense, there can be no guarantee of honest and quality data.

> **I was careful not to romanticise entry and acceptance, because edicts permitting entry can sometimes be a function of power used to serve the interests of narrow groups whose communities are stifled.**

The convenience of the study participants was important to promoting participation (Halcomb et al. 2007). We therefore sought to decide on meeting times that would not disrupt the participants' chores. To mitigate this, we encouraged contact persons to liaise with local gatekeepers and would-be participants and agree on times that were most apposite for them to allow us to administer focus group discussions.

The total sample size was 66 adults (30 females and 36 males) — 62 in six focus groups plus four key informants. As can be seen on Table 2, of the total who participated in focus groups, 28 were females and 34 were males. The mean age was 38.5 years with a range of 54 years. I was not able to gather information on ethnicity, and therefore focus groups were not balanced on that basis. At the time of the evaluation, openly labelling individuals as Hutu or Tutsi risked rekindling violence, even if participants knew among themselves who had belonged to which group. So in this evaluation I sought a process that would facilitate collaboration and unity of purpose and action.

Conducting focus group discussions

I worked closely with my assistants for reasons of data cleaning and to ensure appropriate procedures were followed. At the start of each focus group, session facilitators would welcome participants and stimulate the process with song and dance. This would be followed by individual introductions of all present and an explanation of the purpose of the research. Then they would outline ground rules; in all cases participants were first asked for their consent and assured of the confidentiality of their responses. Participants were also informed that they could withdraw from the focus group session at any time if they so wished without any penalty or loss of any participation rewards. Participants would then sign an informed consent form.

The meetings were held in places and at times that were convenient for all participants. While many authors refer to the importance of neutrality, comfort, and accessibility when

Table 2: Description of Participants and Focus Groups

Place name	Focus group	N	Frequency of gender F = female M = male	Age $\bar{x}$
Ruziba, Kanyosha (Bujumbura)	I	9	F = 5 M = 4	34.7
Ruziba, Kanyosha (Bujumbura)	II	10	F = 5 M = 5	34.6
Buhoro, Itaba (Gitega)	III	9	F = 4 M = 5	43.3
Buhoro, Itaba (Gitega)	IV	15	F = 5 M = 10	33
Ruhororo, Ruhor- oro (Ngozi)	V	10	F = 5 M = 5	44.7
Ruhororo, Ruhor- oro (Ngozi)	VI	9	F = 4 M = 5	44.4
Total		62	F = 28 M = 34	N/A
Average		10.3	F = 4.7 M = 5.7	38.5

choosing a focus group location (Esposito & Powell-Cope 1997; Strickland 1999), the basic rule of thumb is that place should and/or will not inhibit the participants' concentration and answers. For my focus groups, I wanted locations where participants (which included a mix of perpetrators and victims of past violence) were all welcome and would be 'accustom[ed] to gathering and express[ing] themselves freely' (Vissandjee et al. 2002, 830). Locations varied between classroom and community hall. The decision to mix the groups was important to ascertain the extent of CU programming on reconciliation and cohesion, and to understand the nuances transforming the conflict. Separating the groups might have had a hardening effect by replaying the cleavages that CU seeks to address.

To set the tone for the meetings, I ensured we arrived early at each of the focus group session locations. This had important implications for the conduct of the sessions. It allowed us to organise the space, create a welcoming atmosphere, speak to onlookers and participants, and finally to welcome the participants to the venue as they trickled in. Seating arrangements affect group dynamics (Vissandjee et al. 2002, 831). To get the most out of group interaction, I organised all seating in a circle so that all participants could see one another, including the moderator, assistant facilitator, and myself as participant observer.

Payment represents a form of appreciation of and compensation for participation, often offered as food or monetary rewards for those engaged in the research activity. Many researchers assert that important environmental factors in focus groups' conduct include transportation to the session, child-care, food, presence of onlookers or other distractions, and financial remuneration for participation (Halcomb et al. 2007; Vissandjee et al. 2002; Willgerodt 2003). My network partners in Burundi had indicated that many individuals and organisations had been conducting research of various sorts

since the democratic elections of 2005, setting precedents in which participants were offered monetary rewards. I found myself in a quandary with significant potential ramifications for the quality and integrity of the evaluation findings. Although I harboured a fear of the commodification of the process and its likely impact on data, I intended to triangulate focus group findings with anecdotal data and key informant interviews. After carefully considering which option would produce most good and do least harm (the utilitarian perspective), their respective impacts upon the rights of community members (the rights-based perspective), and the best interests of the community as a whole (the common good perspective), I decided to pay participants 1,000 Burundian francs each. This amount equates to roughly US$0.70, which is about the average daily income in Burundi.

Composition of the groups

Tonkiss points out that the 'standard argument within focus group methodology is that group members should be homogeneous in respect of the relevant selection criteria, but unknown to each other' (Tonkiss 2004, 201). Discussing homogeneity, Carey (1994, 229) asserts that 'focus groups should be homogeneous in terms of age, status, class, occupation and other characteristics, as they will influence whether participants interact with each other … also they should be strangers'. Still, there are numerous examples of focus groups in which participants know each other already. Conducting focus groups with participants who do not know each other is not easy to do, especially in small, relatively isolated communities (Strickland 1999). In fact, 'focus groups with people who already know each other and share a sense of common social identity have different strengths and weaknesses from research with groups of comparative strangers' (Holbrook & Jackson 1996, 141). According to Fern (1982) slight differences exist between homogeneous and non-homogeneous groups. The purpose of the focus group should dictate the degree of homogeneity. Heterogeneous groups may be used to produce rich data, whereas homogeneous groups may be used to facilitate rapport (Calder 1977). Discarding the use of homogeneity and strangers in focus groups, Powell and colleagues (1996) allude to pragmatic and methodological motivations including limitations of time and available resources. They suggest that preformed groups, sometimes composed of friends, can offer an encouraging setting constructive to open discussion and, thus, trustworthy research.

Our focus groups were homogeneous in terms of the cultural context, experience of the violent past, and being directly targeted (or not) by the CU programme. There was no further attempt to achieve homogeneity, as all groups were mixed in terms of age, gender, and ethnicity. For the most part, all group participants knew each other very well, whether as relatives, friends, colleagues, or acquaintances. The standard argument that participants should be strangers did not hold for my focus groups. There were several reasons for this:

1. In Burundi, societal dynamics have been affected by several decades of intermittent and cyclic ethnic-induced violence and so made people, especially in rural communities, know just about everybody as a means of resilience.
2. It would have been difficult to recruit an appropriate sample of adult Burundi women and men had the main criteria been that they should not know each other, and, regardless of sample population, it may well have been impossible to avoid people knowing each other.
3. Given the nature of my study (not to mention the ex post facto observations made in the course of the focus group discussions), participants might have been less

willing to speak about personal matters concerning their past violent experiences and coping strategies had they not known each other. By contrast, when participants already knew each other, dialogue was often more interactive and others apparently were comfortable being at variance with particular points raised.

Group dynamics

An important feature of focus groups is the interaction among participants that is expected in order to elicit rich experiential data (Webb & Kevern 2001). This interaction, sometimes called 'group dynamics' (or group forces), is an integral part of the focus group process in which participants engage in dialogue with each other, rather than directing the debate to the moderator facilitating the discussions. The significance of group dynamics is in the assumption that the interaction will be creative by widening the range of responses, triggering forgotten details of experiences, and overcoming introversion that may otherwise discourage participants from disclosing information (Merton et al. [1956] 1990).

In our focus groups and at the beginning of each session, we tried hard to create an atmosphere of trust. We endeavoured to convey that participants' contributions were important and that we were not there to assess the 'wrongs' or 'rights' of their responses. This dynamic was important for me as principal researcher for two reasons. First, it helped to identify the conditions that promoted interaction and open discussion of the participants' views and experiences of the peacebuilding work done by non-governmental organisations in Burundi. Second, creating an atmosphere of confidence to allow unrestricted dialogue assisted in the analysis of data by affording an understanding of what was happening in the group at the time a particular piece of information was mentioned, as well as why it might have been happening. I found this to be in harmony with Krueger (1994) and Kitzinger (1994), who both identify the substance of interface in focus groups, but for rather dissimilar reasons. For Krueger, interaction between participants is a helpful device to encourage discussion on a topic. It performs a useful instrumental function in gathering data. Kitzinger, on the other hand, identifies a more central role for interaction, as the central analytical resource; it is intrinsically valuable, not simply an efficient way of gathering data. As a result, Kitzinger underscores the importance of concentrating on interaction between participants in analysis, to the point of making it a defining feature of focus group method.

Creating an atmosphere of confidence to allow unrestricted dialogue assisted in the analysis of data by affording an understanding of what was happening in the group at the time a particular piece of information was mentioned, as well as why it might have been happening.

Several unanticipated factors affected group dynamics. In Burundi, social events typically attract the community at large. Also, white 4 × 4 Toyota vehicles are associated more often than not with non-governmental organisations and with hand-outs of some variety. Other villagers who heard of our visit were present on the days we conducted the focus groups just out of curiosity. These bystanders inadvertently created pockets of outside disturbance of the focus groups. To mitigate distractions posed by gatherings of bystanders, the research team usually communicated with them to explain the nature and purpose of our visit.

Internal factors concerned the factors of collective isolation, power relations, and self-censorship and conformity. Focus groups are essentially social moments in the sense that

they are conducted with a uniquely composed group of participants at a particular point in time. These focus groups were not natural situations, even though they might have been social congregations. Some participants were clearly tense and uncomfortable because the participants, as argued by Jowett, 'were assembled together at my request to reflect on particular issues which had been set by [me]' (Jowett & O'Toole 2006, 458). Intermittent silences and unmistakable discomfort, especially in focus groups I and III, suggested that participants found themselves in an artificial and undesirable setting. This formed a kind of collective isolation. Accordingly, 'it would be naïve to assume that group data is by definition "natural" in the sense that it would have occurred without the group being convened for this purpose [peacebuilding]' (Kitzinger 1994, 106).

There is a general assumption that the researcher wields power over the researched. Importantly, my experience in one encounter supported the argument that strength in numbers can profitably disrupt power relationships in focus group discussions. I must mention that I am a black Zimbabwean male living in South Africa, and was introduced thus in all opening stages of focus group discussions. In focus groups II and VI, when participants were asked what opinion and advice they would suggest to secondary stakeholders like the state and civil society organisations to prevent future violence and build peaceful communities, participants turned that question to me. I realised later that my background had made them increasingly confident to ask me about my experience and my own community back in Zimbabwe, as we worked to mitigate the violence that had turned the country from the 'bread basket of Southern Africa' into a 'basket case'. I was astonished at how my background eroded the traditional barriers associated with the research relationship. I began to think about just how deep the sense of grassroots alienation from the peacebuilding process went, and how this might impact on the overall efforts to build durable peace. I began also to piece together how sometimes Track II advocates and bottom-up approaches are disparaged in favour of top-down and Track I actors.

> **Strength in numbers can profitably disrupt power relationships in focus group discussions.**

One worrying phenomenon I witnessed in group interactions centred on the issue of self-censorship and conformity. In some instances I felt some participants' responses were influenced by others' behaviour. This was most evident in focus groups I and IV. The major pitfall here, as Carey suggests, was the risk that a participant would

> tailor his or her contributions to be in line with perceptions of the group members and/or the leader [and] withhold potential contributions, often due to a lack of trust of the leader, or other members, [as well as] mentally reconstruct their experiences on the basis of the on-going dialogue. (Carey 1994, 236)

Conformity with the group and/or conceivably with the prompts of the moderator was notable in sessions I and IV. Some participants were a little nervous of speaking in the group situation. Maybe this also could have been generated by my presence as a visitor. The groups knew my assistants better, as they had worked together in the course of CU programming. As I mentioned earlier, all groups were heterogeneous in terms of gender. In focus groups I and IV, I found that women tended to speak first, which was a contrast to O'Toole's experience (Jowett & O'Toole 2006) and the generally widely held belief in traditional African society that women do not speak in public and in the

> **In these heterogeneous focus groups, women were generally more likely to be vocal in their disagreements with others, and generally dominated the discussions.**

presence of men. In these heterogeneous focus groups, women were generally more likely to be vocal in their disagreements with others, and generally dominated the discussions. Men took much longer to join in the dialogue and, when they did contribute, tended only to support and echo each other.

Each time we began a session, I harboured three fears. First, the heterogeneous nature of focus group composition (women and men; Hutu, Tutsi, and Twa; former rebels, demobilised soldiers, and civilian victims of the violent past) led me to have concerns that the focus group might not have been the most suitable method for my study. I worried that the group mix might create a recipe for self-censoring by participants who didn't feel strong enough to voice their opinions in a space shared with one who hurt them, and/or vice versa. Second, I worried that the dialogue would be pretentious and lack real substance. Third, I feared that the discussions might be affected by an overriding desire to please me, rather like pupils when they attempt to guess what a teacher wants to hear and volunteer it first (Grugeon & Woods 1990). However, my fears were allayed once the sessions gained momentum. Participants were loquacious and steered the debate to many bearings, epitomizing what has been called 'the synergistic trajectory of the focus group method' (Stewart & Shamdasani 1990, 16).

Recording, analysing, and interpreting data

Recording the proceedings of the focus groups' conduct is important because it allows opportunities for rigorous analysis of the data to enhance comprehension and appreciation of the issues that are being researched. We took time to explain clearly that we needed both to record and to transcribe the focus groups, because the verbatim language had to be deciphered from the local Kirundi into English. When permission was obtained from participants in each focus group session, we tape-recorded the focus group proceedings. They also gave consent to photography. While the chief moderator for a session facilitated the discussions, the assistant moderator made written notes. As an observer, I made notes of the non-verbal interaction and also made notes of the frequency of participant interaction. I did this by coding participants according to how they sat in the room. Then, using the tally method, I indicated every time any one participant spoke. I did this to determine the dominant and timid or silent participants. An indication of the origin of each idea and/or statement provided an opportunity to cross-link concepts and themes with roles and identities and personal attributes. Also, to keep track of the new knowledge gathered, immediately after each focus group session I took time with my assistants to briefly review the data and notes gathered. This was crucial because it helped me to clarify notes and add comments about things that might have eluded the team during the focus group process.

The analysis and interpretation of focus group data can be very complex (De Vos et al. 2007). To analyse these focus group data, I employed the content analysis approach, also called 'text' or 'narrative data analysis'. The study generated stories (narratives) from focus group interviews, in-depth interviews, and observations including the group dynamics of both verbal and non-verbal interactions. Basically, it involved a three-stage process of data reduction, display, and interpretation and corroboration, blending qualitative and quantitative elements. The process was fluid, drifting and looping back and forth between the steps to refine indexing, re-reading texts, and revising other aspects of the analysis.

Finally, I brought all the data together by attaching meaning and significance to the analysis. I did this by eliciting and sketching key points and important findings from data display and categorized data. To allay the problem of insufficient interpretation, I spent

time and consulted with CU staff and other key informants to synthesize the meanings of the findings. I did this procedure guided by the following questions: What are the key lessons? What new things did I learn? What will those who use the results of the evaluation of CU be most interested in knowing? What has application to other programmes, settings, and studies?

Conclusion

Given the foregoing discussion, not all violently divided contexts are likely to generate unreliable data for the monitoring and evaluation of peacebuilding work.

Thinking creatively about the methods and procedures used in field evaluation is crucial to one's ability to collect quality data, analyse the data, and ascertain the peacebuilding and development impact on individual and community relationships and infrastructure.

Thinking creatively about the methods and procedures used in field evaluation is crucial to one's ability to collect quality data, analyse the data, and ascertain the peacebuilding and development impact on individual and community relationships and infrastructure. Meticulous discussions of CU's impact derived from this field study.[1]

Indeed, ethical concerns of field research in unstable contexts are often puzzling. The complexity of building rapport has immense ethical challenges for the researcher and those researched. For instance, paying participants is rooted in the knowledge of a particular idiosyncratic place and time. If active violence were ongoing at the time of evaluation, the amount that it would have been acceptable to pay participants for their time might have been higher (to serve as a sort of humanitarian aid when livelihoods were otherwise disrupted) or lower (in the event that the payments brought unwanted attention to participants from predatory actors). Also, the homogeneity versus heterogeneity question might be informed by such considerations, as pre-conflict Burundian society was presumably much more sensitive to ethnically heterogeneous groups than a post-conflict society in which such distinctions are being downplayed.

In closing, to put it most simply, ethical considerations that produce the most good and do little harm (utilitarian approach), that are respectful of the rights of community members (rights approach), without discrimination (justice and fairness), and serving whole-community interests not just individual interests (common good approach) are critical to understanding the how, what, and why of the best and promising peacebuilding and development programming.

SYLVESTER BONGANI MAPHOSA is the chief research specialist at Africa Institute of South Africa (AISA). He holds a PhD in conflict resolution and peace studies, an MA in peace and governance, a BEd in environmental sciences, and a diploma in education. He has extensive field experience in research, government, higher education, non-governmental organisations, and capacity-building for peace.

Endnote

[1] Maphosa 2009a; 2010. See also Maphosa 2009b. The latter is available on request from the Director of CU, Father Emmanuel Ntakarutimana: +257 22 245045.

References

Anderson, M. & **Olson, L.** 2003, *Confronting War: Critical Lessons for Peace Practitioners*, Cambridge, MA: Collaborative for Development Action.

Babbie, E. & **Mouton, J.** 2007, *The Practice of Social Research*, Cape Town: Oxford University Press Southern Africa.

Calder, B.J. 1977, 'Focus Groups and the Nature of Qualitative Marketing Research' in *Journal of Marketing Research* 14: 353–364.

Carey, M. 1994, 'The Group Effect in Focus Groups: Planning, Implementing, and Interpreting Focus Group Research' in Morse, J. eds, *Critical Issues in Qualitative Research Methods*, London: Sage.

Church, C. & **Shouldice, J.** 2002, *The Evaluation of Conflict Resolution Interventions: Framing the State of Play*, Londonderry: INCORE.

Church, C. & **Shouldice, J.** 2003, *The Evaluation of Conflict Resolution Interventions, Part II: Emerging Theory and Practice*, Londonderry: INCORE.

D'Estree, T., **Fast, L.A.**, **Weiss, J.N.** & **Jakobsen, M**. 2001, 'Changing the Debate about "Success" in Conflict Resolution Efforts' in *Negotiation Journal* 20: 2: 101–113.

De Vos, A., **Strydom, H.**, **Fouche, C.** & **Delport, R.** 2007, *Research at Grassroots for the Social Sciences and Human Services Professions*, Pretoria: Van Schaik.

Douma, N. & **Klem, B.** 2004, *Civil War and Civil Peace. A Literature Review of the Dynamics and Dilemmas of Peacebuilding through Civil Society*, report Netherlands: Institute of International Relations.

Esposito, N. & **Powell-Cope, G.** 1997, 'Avoiding Pitfalls in Focus Group Research' in *Science Nursing* 14: 1: 31–33.

Fern, E.F. 1982, 'The Use of Focus Groups for Idea Generation: The Effects of Group Size, Acquaintanceship, and Moderator on Response Quantity and Quality' in *Journal of Marketing Research* 19: 1–13.

Grugeon, E. & **Woods, P.** 1990, *Educating All: Multicultural Perspectives in the Primary School*, New York: Routledge.

Halcomb, E.J., **Gholizadeh, L.**, **DiGiacomo, M.**, **Phillips, J.** & **Davidson, P.M.** 2007, 'Considerations in Undertaking Focus Group Research with Culturally and Linguistically Diverse Groups' in *Journal of Clinical Nursing* 16: 1000–1011.

Holbrook, B. & **Jackson, P.** 1996, 'Shopping Around: Focus Group Research in North London' in *Area* 28: 2: 136–142.

Johnson, R.B. & **Onwuegbuzie, A.J.** 2004, 'Mixed Methods Research: A Research Paradigm Whose Time Has Come' in *Educational Researcher* 33: 7: 14–26.

Johnson, R.B., **Onwuegbuzie, A.J.** & **Turner, L.A.** 2007, 'Toward a Definition of Mixed Methods Research' in *Journal of Mixed Methods Research* 1: 2: 112–133.

Jowett, M. & **O'Toole, G.** 2006, 'Focusing Researcher's Minds: Contrasting Experiences of Using Focus Groups in Feminist Qualitative Research' in *Qualitative Research* 6: 4: 453–472.

Kitzinger, J. 1994, 'The Methodology of Focus Groups: The Importance of Interaction between Research Participants' in *Sociology of Health and Illness* 16: 1: 103–121.

Kreuger, R.A. 1994, *Focus Groups: A Practical Guide for Applied Research*, 2nd edn. Thousand Oaks, CA: Sage.

Kriesberg, L. 2007, *Constructive Conflicts*, Lanham, MD: Rowman, Littlefield.

Kuzel, A.J. 1992, 'Sampling in Qualitative Inquiry' in Miller, B.F. & Miller, W.L. eds, *Doing Qualitative Research*, Newbury Park, CA: Sage.

Lederach, J.P. 1997, *Building Peace: Sustainable Reconciliation in Divided Societies*, Washington: US Institute for Peace.

Lieberfeld, D. 2002, 'Evaluating the Contributions of Track Two Diplomacy to Conflict Termination in South Africa, 1984–90' in *Journal of Peace Research* 39: 3: 355–372.

Magill, R.S. 1993, 'Focus Groups, Program Evaluation, and the Poor' in *Journal of Sociology and Social Welfare* 20: 1: 103–114.

Maphosa, S.B. 2009a, 'Building Peace from Below: The Centre Ubuntu in Burundi' in *Africa Peace and Conflict Journal* 2: 2: 58–71.

Maphosa, S.B. 2009b, *A Review of CENTRE UBUNTU Peace Building Work: Burundi, Burundi*, Evaluation Report Bujumbura: Centre Ubuntu.

Maphosa, S.B. 2010, 'Analysing Programme Strategies Using the Reflective Peace Practice Matrix and the Seven Functions of Civil Society in Peace Building: Reflections on a Case Study from Burundi' in *GhandiMarg Journal* in 31: 4: 711–736.

Maynard-Tucker, G. 2000, 'Conducting Focus Groups in Developing Countries: Skill Training for Local Bilingual Facilitators' in *Qualitative Health Research* 10: 3: 396–410.

Mays, N. & Pope, C. 1995, 'Rigour and Qualitative Research' in *British Medical Journal* 311: 109–112.

Menkhaus, K. 2004, *Impact Assessment in post-conflict Peacebuilding: Challenges and Future Directions*, International Peacebuilding Alliance, Interpeace: http://pdf2.hegoa.efaber.net/entry/content/643/3_Impact_Assessment_in.pdf, accessed January 9, 2009.

Merton, R.K., Fiske, M. & Kendall, P.L. [1956] 1990, *The Focused Interview: A Manual of Problems and Procedures*, New York: Free Press, 2nd edn.

Mitchell, C. 1993, 'Problem-Solving Exercises and Theories of Conflict Resolution' in Sandole, D. & Merwe, H. eds, *Conflict Resolution: Theory and Practice*, Manchester: Manchester University Press.

Paffenholz, T. & Reychler, L. 2007, *Aid for Peace: A Guide to Planning and Evaluation for Conflict Zones*, Baden-Baden: Nomos.

Powell, R.A., Single, H.M. & Lloyd, K.R. 1996, 'Focus Groups in Mental Health Research: Enhancing the Validity of User and Provider Questionnaires' in *International Journal of Social Psychology* 42: 3: 193–206.

Ross, M. 2004, 'Conceptualizing Success in Conflict Resolution Interventions' in *Peace and Conflict Studies* 11: 1–18.

Rouhana, N. 2000, 'Interactive Conflict Resolution: Issues in Theory, Methodology, and Evaluation' in Stern, P. & Druckman, D. eds, *International Conflict Resolution after the Cold War*, Washington: National Academy Press.

Strickland, C.J. 1999, 'Conducting Focus Groups Cross-Culturally: Experiences with Pacific Northwest Indian People' in *Public Health Nursing* 16: 3: 190–197.

Stewart, D.W. & Shamdasani, P.N. 1990, *Focus groups: Theory and practice*, Newbury Park, CA: Sage.

Tonkiss, F. 2004, 'Using Focus Groups' in Seale, C. eds, *Researching Society and Culture*, London: Sage.

Vissandjee, B., Abdool, S.N. & Dupure, S. 2002, 'Focus Groups in Rural Gujarat, India: A Modified Approach' in *Qualitative Health Research* 12: 6: 826–843.

Webb, C. & Kevern, J. 2001, 'Focus Groups as a Research Method: A Critique of Some Aspects of Their Use in Nursing Research' in *Journal of Advanced Nursing* 33: 6: 798–805.

Wilkinson, S. 1999, 'How Useful Are Focus Groups in Feminist Research?' in Barbour, R.S. & Kitzinger, J. eds, *Developing Focus Group Research: Politics, Theory and Practice*, London: Sage.

Willgerodt, M.A. 2003, 'Using Focus Groups to Develop Culturally Relevant Instruments' in *Western Journal of Nursing Research* 25: 7: 798–814.

Guiding Principles for Donors to Foster Better Peacebuilding Evaluation: An Update on the Consultation Process

Andrew Blum & Melanie Kawano-Chiu
Published online: 19 Sep 2013.

PLEASE SCROLL DOWN FOR ARTICLE

BRIEFINGS

GUIDING PRINCIPLES FOR DONORS TO FOSTER BETTER PEACEBUILDING EVALUATION: AN UPDATE ON THE CONSULTATION PROCESS

ANDREW BLUM AND
MELANIE KAWANO-CHIU

Introduction

For the past two years, the Alliance for Peacebuilding (AfP) and the United States Institute of Peace (USIP) have overseen the Peacebuilding Evaluation Project (PEP). The project was among the first to convene peacebuilding funders, implementers, and analysts to identify concrete strategies for improving how the peacebuilding field assesses programming and communicates its programme results.[1] The core premise of the PEP is that collective approaches, particularly those which bring donors and implementers together, are necessary to genuinely improve evaluation within the peacebuilding field.[2] Individual organisations can always make improvements in their practice, but this must be combined with collective approaches if key challenges are to be overcome.

The Guiding Principles was one of the collective approaches that PEP participants suggested. The goal of the Guiding Principles is to create a common set of standards regarding how the funder–

implementer relationship should be structured in regard to evaluation practice.[3] As is the case with all standards-making processes, the purpose of the standards is to codify good practice, create common expectations within the field, and create a common framework within which more productive communication can take place.[4]

The Consultations

Beginning in March 2012, AfP and USIP began a consultative process to develop a draft set of 'Guiding Principles for Peacebuilding Donors' based on initial discussions from PEP. As part of this process, small-group feedback sessions were held in six peacebuilding hubs: Washington (13 September, at USIP), New York (3 October, at UN Peacebuilding Support Office), Geneva (11 October, at Geneva Peacebuilding Platform), London (12 October at the United Kingdom Department for International Development), San Francisco (16 October, at Humanity United), and Nairobi (13 November, at World Bank, Global Center on Conflict Security and Development).

The six small group meetings ranged in size from seven to twenty participants. The make-up of the group at each meeting varied, but participants included evaluation experts and representatives from donor organisations, multilateral organisations, and implementing organisations working on peacebuilding.[5] For each consultation, participants were provided a draft set of Guiding Principles ahead of time, developed in consultation with participants of the Peacebuilding Evaluation Project. The draft document contained five sections: Purposes of Evaluation; Resources and Capacity-Building; Methodology; Use, Sharing, and Learning; and Existing Good Practices. The participants were led through a discussion to solicit feedback on (1) each section of the document; (2) what was missing from the document; (3) strategies for disseminating

105

and getting buy-in for the document, as well as other next steps.

To provide a clearer idea of the nature of these principles, below are two examples from the draft document:

- *From the purposes section:* Recognise that evaluations should not be confined to a summative event at the end of a programme or just for external stakeholders. Encourage evaluative thinking throughout the project cycle.

- *From the methodology section:* Discuss and agree upon definition and standards for credible evidence prior to the beginning of the project.

Problems, Solutions, Debates

Four key debates arose during the consultation process. These debates provide a useful means both to summarise the consultations and to illuminate key issues in regard to the Guiding Principles and peacebuilding evaluation more broadly.[6] Clearly, there is still an open and healthy debate on many of the issues that the Guiding Principles seek to address, and this should continue. Below are summaries of the four key debates.

1. Should this document apply just to donors or to donors and implementers?

There was debate at the consultations regarding whether the document should serve as a set of general guidelines on evaluation that would apply to both donors and implementers, or should specifically target donor practice. A frequent and related question often asked concerned the relationship of the document to other key documents in the peacebuilding field, particularly the new Organisation for Economic Cooperation and Development's (OECD) Development Assistance Committee (DAC) Guidelines on Evaluating Peacebuilding Programming (OECD 2012).

The different approaches represent solutions to different problems. A set of general guidelines for both donors and implementers would primarily be an educational exercise. It would provide information to everyone on what is considered good practice in regard to peacebuilding evaluation. In our view, this is one of the key contributions of the OECD guidelines. However, the Guiding Principles would focus specifically on basic principles that can underlie donor practice, while the OECD guidelines focus more broadly on good practice for conducting evaluations.

In contrast, a document that specifically targets donor practice is designed to address the power imbalance between donors and implementers which is inherent to the funding relationship. While donors are accountable to their ultimate funders, often legislatures, there is little downward accountability from the donor to the implementer. There is very rarely language in funding documents describing what the donor is responsible for, how results will be used, and so on. The fact that there are few standards to which donors are held accountable contributes to poor evaluation practice. For instance, several participants noted that donors often change evaluation requirements in the middle of a project if staff or organisational priorities change. Other implementers noted that they received no information regarding how reports sent to the donor are used. Instances of poor practice such as these are not just a recipe for frustration, but create among implementers a souring to honest evaluation, leading in many cases to a retreat to seeing evaluation as a box-checking exercise to assuage donors.

A clear, consistent set of principles that describe what is expected of donors in regard to supporting evaluation would allow donors and implementers to hold

themselves mutually accountable for good evaluation practice. This would improve evaluation practice directly, through better donor practice, and indirectly, through implementers having increased confidence that good evaluation practice will be rewarded by donors.

2. How different is the challenge of peacebuilding evaluation compared with evaluation in other sectors?

There is a genuine disagreement within the peacebuilding field regarding whether evaluation of peacebuilding programming is uniquely difficult, that is, more difficult than evaluation in other international development sectors such as health, education, livelihood programming, or democracy and governance. Some participants argued that the complexity of peacebuilding programming and the long timeframe and the non-linearity of the social changes required to build peace require unique evaluation methodologies. These participants indicated that they saw the principles as too generic and wanted them to relate to peacebuilding programming more specifically. These participants saw the principles as an opportunity to push back against the 'measurement fetish' among donors, and to articulate a different approach to peacebuilding evaluation more in line with the challenges the sector poses, although it was not clear from the conversations what exactly these alternatives would look like.

In contrast, other participants argued that evaluation of peacebuilding programming is challenging, but not uniquely so and no more difficult than evaluation in other development sectors. They see the principles as an opportunity to bring evaluation in the peacebuilding field up to the level that other sectors have reached. As a result, they were less concerned with the lack of peacebuilding-specific language within the Guiding Principles.

This disagreement is not an easy one to resolve and to some extent reflects peace-

building evaluation's current stage of development. The final Guiding Principles will need at the very least to provide guidance to donors regarding the need to clearly and transparently communicate their views on the question of how uniquely challenging peacebuilding evaluation is.

3. How aspirational should the document be?

There was a debate in the consultation sessions regarding whether the Guiding Principles should articulate principles for the field to aspire to, or be limited to a clear statement of good practice given current realities. This argument also represented a difference regarding what type of change the initiative is seeking to bring about. Those arguing for a clear statement of current practice are asking donors to adhere to already existing good practice. Donors, for instance, already commit in numerous ways to transparency. The Guiding Principles would provide a form of social pressure on donors to live up to standards they have already set for themselves.

In contrast, those arguing for a more aspirational document see the Guiding Principles as an opportunity to articulate new practices that donors could undertake to improve evaluation in a more transformational way. For instance, many participants argued that donors should move away from a sole focus on project-level evaluations. Donors are most comfortable funding projects and then seeing the evaluation of those projects either by the organisation itself or by an external evaluator. This funding model creates several funding gaps. For instance, many evaluation-related common goods go unfunded. Participants mentioned, for instance, that there is virtually no one seeking to aggregate evaluation results across multiple projects as a means of understanding how several projects worked together (or not) to create broader peacebuilding outcomes in a given context. In addition, there is not sufficient funding for shared repositories of evaluation

resources (surveys, indicator menus, expert databases, etc.). Finally, funding for evaluation capacity within organisations suffers because this capacity is an organisational capacity that is not tied to a specific project.

However, shifting from a project-level funding model is aspirational. It is something that the field could articulate as an important goal, and is actually something that is starting to happen in other development sectors where there is an increasing focus on programme-level and sector-level evaluation. However, it is not something that donors could, or should, be held accountable for at the moment, given the systemic changes that would have to take place before project-level funding could be de-emphasised in relation to other models of funding.

4. How much authority should the document have behind it? Should we seek to have the Guiding Principles adopted as the official policy of donors?

There was debate at the consultations between those who argued for a more substantive list of principles that would guide donor policy and practice in a more comprehensive way, and those who argued for pithy, memorable principles that would establish basic expectations. Those who argued for more substantive principles envisioned the Guiding Principles as a more authoritative document, whose principles, for instance, might be included in the policy documents of donors. Those arguing for a briefer list envisioned a more open, voluntary form of endorsement by donors.

Again, these two approaches see the document serving different purposes. Those arguing for more substantive guidelines want the document to improve practice through directly changing the formal policies and practices of donors in an authoritative way. Those arguing for a more concise list saw the principles as

altering practice by improving communication between donors and implementers. Participants noted that it is often difficult for individual implementers to engage in honest communication with their donors. As a result, donors and implementers struggle to establish shared understandings of all aspects of evaluations, from their initial purpose to the eventual use of results. Thus, participants who argued for a shorter list of principles saw the document as creating change by providing a basic set of guideposts that would improve practice through improving communication and establishing shared understanding between donors and implementers.

Conclusion

As is normally the case, the consultations did not provide answers, but instead successfully clarified the key issues to be resolved. Each of the debates discussed above illuminates subtle, nuanced, and sophisticated issues regarding the types of change that the Guiding Principles seek to create, and the ways those changes may best be achieved. During the next stage of the initiative, additional feedback on a revised Guiding Principles document will be solicited from a wide range of donors and implementers. Given the issues raised by the debates discussed, decisions will need to be made regarding how the Guiding Principles can be structured to create the most positive change.

As the final shape of the Guiding Principles document comes together, efforts will be made to begin to gain commitments from donors to work towards upholding the Guiding Principles in their practice.[7] In the longer term, regardless of how the debates described in this briefing are resolved, the Guiding Principles will need to receive broad endorsement from throughout the peacebuilding field if they are to have any impact. Toward this end, AfP, through its policymaker engagement programme 3P Human Security, is working under the auspices of the newly formed Peacebuilding Evaluation Consortium to

foster widespread support for the Guiding Principles within US government agencies and among other bilateral and multilateral donors.

DR ANDREW BLUM is currently Vice-President, Program Management and Evaluation at USIP. He is responsible for improving USIP's programming through-out the project cycle from design to implementation to evaluation and learning. Andrew received a PhD from the University of Southern California and a BA from the University of Virginia.

MELANIE KAWANO-CHIU is the Program Director at AfP, where she manages programmes on evaluation, systems, communications, and prevention. Previously, she managed the BEFORE Project, taught at a university in Vietnam, and worked at the International Foundation for Electoral Systems. Melanie has a BA from Scripps College and an MA from the University of Denver's Joseph Korbel Graduate School.

Endnotes

[1] For a list of PEP participants, see Kawano-Chiu (2011, 32).

[2] See Blum 2011.

[3] The Good Humanitarian Donorship Principles are a similar undertaking: http://www. goodhumanitariandonorship.org/gns/princip les-good-practice-ghd/overview.aspx, accessed 14 December 2012.

[4] For another example of a standards-making process, see the Good Humanitarian Donorship Principles: http://www.goodhumanitaria ndonorship.org/gns/principles-good-pra ctice-ghd/overview.aspx, accessed 30 April 2013.

[5] The meetings were held under the Chatham House Rule, so there will be no attribution in this briefing. The content of this briefing are the responsibility of the authors alone.

[6] It's important to emphasise that the role of this briefing, as a report on the consultations, is to summarise these debates, not to adjudicate them or to articulate the preferences of the authors.

[7] For more information on the process or to provide feedback on the process, please contact either of the authors.

References

Blum, A. 2011, *Improving Peacebuilding Evaluation: A Whole-of-Field Approach*, Washington, DC: USIP.

Kawano-Chiu, M. 2012, *Starting on the Same Page: A Lessons Report from the Peacebuilding Evaluation Project*, Washington, DC: AfP.

OECD. 2012, 'Evaluating Peacebuilding Activities in Setting of Conflict and Fragility: Improving Learning for Results', DAC Guidelines and Reference Series.

Developing Oecd Dac Guidance on Evaluating in Settings of Violent Conflict and Fragility

Megan Grace Kennedy-Chouane
Published online: 19 Sep 2013.

PLEASE SCROLL DOWN FOR ARTICLE

DEVELOPING OECD DAC GUIDANCE ON EVALUATING IN SETTINGS OF VIOLENT CONFLICT AND FRAGILITY

MEGAN GRACE
KENNEDY-CHOUANE

Introduction

Recognising the need for improved learning and accountability in settings of violent conflict and state fragility, the Organisation for Economic Cooperation and Development's (OECD's) Development Assistance Committee (DAC) launched an initiative to develop evaluation guidance in 2004/2005. Two of the DAC's specialised networks, working on evaluation and on conflict, joined forces to take this work forward. Draft guidance was published in 2007 and used for evaluations in different conflict settings. Drawing on these practical experiences, the guidance was further developed and finalised in 2012. It aims to strengthen evaluation practice and support the community of donors, country partners and implementing organisations to improve the results of peacebuilding support. This article looks at the process of developing the guidance and explores some of the lessons emerging from that experience. It highlights the collaborative nature of the process, which presented both challenges and strengths, as well as the value of an extended pilot phase.

The Need for Better Evaluation in Conflict Settings

The impetus for developing guidance came from donors in the Conflict Prevention and Development Co-operation (CPDC) network, a sub-committee of the DAC made up of policy experts working on conflict in the context of development cooperation. This group identified a number of weaknesses around learning and accountability in the diverse fields of development cooperation related to peace and conflict. There was widespread agreement that, despite new clarity on the links between conflict and poverty reduction and guidance on how donors could support peace, the analysis of implementation and results remained weak and there was little systematic learning from experience. While early work of the DAC CPDC recognised that 'effective systems monitoring and evaluation can help ensure positive results', the actual implementation of evaluative analysis seemed to be lagging.

Results in countries affected by violence and state fragility were also considered disappointing. Particularly as spending was scaled up in the post-2001 period, results were seen as not commensurate with the overall investment of international and national stakeholders. In 2004, the Norwegian Agency for Development Cooperation commissioned the Joint Utstein Peacebuilding study (Smith 2004). The report looked at 336 peacebuilding projects and found that a major strategic deficit existed between the articulation of policy and efforts to translate this policy into practice. For example, more than half of the projects identified did not show any link to a broader strategy for the country in which they were implemented. The report highlighted the dearth of rigorous evaluations in the fields of peace and conflict prevention.

110

Shortly after the release of the Utstein study, the CPDC, which had also been working on evaluation, approached the DAC's Network on Development Evaluation (EVALNET) for support with tackling evaluation challenges, and collaboration between the two communities began. The group had identified a number of evaluation challenges — notably around selecting suitable indicators, dealing with attribution, understanding complexity and linking smaller projects with broader peacebuilding processes. Coupled with these concerns, various practical and methodological challenges make evaluating these types of programmes difficult. Data are scarce, objectives frequently ill-defined and the logic underpinning interventions often murky — to name just a few of the barriers to evaluation in conflict settings.[1] A first meeting was held in Oslo in 2005, where it was decided to develop DAC guidance.

This collaborative initiative hoped to tackle some of the methodological issues facing evaluators, who were increasingly called to assess the effectiveness and results of conflict-related interventions. It was also an attempt to address the broader strategic gap identified in conflict prevention and peacebuilding (Smith 2004; OECD & CDA 2007). It was hoped that better understanding and use of evaluation would provide much-needed hard evidence about the role of development cooperation in transforming conflict and enabling development in settings of state fragility.

A Collaborative Approach to Improving Practice

The cross-community dimension of this work has been a unique strength, and perhaps also the largest challenge, of the process of developing the guidance. The guidance is the result of several years of collaboration between the peacebuilding and evaluation communities, primarily within the two DAC networks EVALNET

and CPDC. The guidance targeted diverse stakeholders: supporting improved evaluation knowledge among those working on conflict and peacebuilding, and helping evaluators better understand the particularities of working in these fields.

Early on in the process it became clear that there was a need for greater clarity of concepts as evaluators asked members of the CPDC, 'What is it that you want to evaluate?' Over the course of several workshops in 2005–2008 representatives of development agencies and foreign ministries, outside experts and developing country partners met to discuss challenges and work towards a consensus on what kinds of interventions would be covered by the guidance. This involved developing a concept of what constituted peacebuilding and conflict prevention and, therefore, what kinds of activities would use the proposed evaluation approach. Much of the confusion reflected the fact that these were relatively new, fast-evolving fields of research, policy and practice, and definitions had not yet been established.

From the outset, the two communities debated whether or not evaluation approaches for conflict prevention and peacebuilding were really 'different' from other types of complex evaluation. Many of the challenges identified as 'unique' by peacebuilding experts (such as missing baseline data or difficulties attributing society-level impacts to individual projects) were very familiar to evaluators — many of whom argued that the CPDC should simply use existing tools. In the end, it was generally accepted that there were enough particularities to warrant a specialised approach. This decision reflected what would become a core position of the International Network on Conflict and Fragility (INCAF — the successor to the CPDC), namely, that fragile states require a specialised approach (OECD DAC & NORAD 2011). On the part of peacebuilding practitioners, some viewed evaluation as limiting — unable to capture the complexity and

value of their work. And yet they hoped it would provide stronger evidence that what they were doing mattered. The process of developing the guidance was thus not just about drafting a document, but about building bridges with shared understanding to enable both parts of the collaboration to contribute to and benefit from the work.

There was also much debate on how to distinguish between working 'in' and working 'on' conflict, that is, between conflict sensitivity and effectively transforming conflict. In the context of development cooperation this raised the question of how activities with primarily humanitarian or development objectives would be treated, and whether or not they should be judged using DAC evaluation criteria — efficiency, effectiveness, sustainability, relevance and impact — as these relate specifically to broader peace objectives. Clarity emerged on this question during the testing phase, described below.

Shortly after the draft guidance was published, a process began in the DAC to bring together work on conflict prevention and fragile states. CPDC was merged with the Fragile States Group to become the DAC INCAF. This reflected the growing consensus in development circles that statebuilding and peacebuilding processes must be dealt with concurrently. As understandings of good practice in these fields continued to emerge and policy guidance was better established, concern over definitions faded. In the meantime, the group developing operational evaluation guidance reached a consensus to use a working definition and four broad descriptive categories of activities: socio-economic development, good governance, the reform of security and justice institutions, and truth and reconciliation. In the 2012 guidance, even these groupings were dropped in favour of a more practical approach that encouraged use of guidance based on the context and purpose of the evaluation and its specific evaluation questions. This evolution reflects in part

that the evaluation community took a greater role during the application phase, while some of the peacebuilding practitioners, who were primarily concerned with definitional issues, disengaged. While an evaluation guidance could have no doubt been more quickly and easily produced by simply commissioning the work from an expert consultant, the consultative process was critical in producing a final product that was not only of high quality, but also valued and owned by its intended audience. By the end of the process there seemed to be a much wider common ground between the two communities: peacebuilding experts were more open to the potential benefits of rigorous evaluation and evaluators more willing to acknowledge that there were some particularities to evaluating in these fields.

Testing of the Guidance

The OECD published the draft *Guidance on Evaluating Conflict Prevention and Peacebuilding Activities* in 2008 and launched a two-year application phase to try out the approach in real conflict settings. The testing phase aimed to produce policy insights as well as feedback on the guidance, looking in particular at the applicability of the guidance to different types of evaluations, levels (country, programme, project and policy/strategic) and use by different actors and in different conflict contexts. To avoid producing evaluations purely for the sake of testing the guidance, pilots were not centrally managed and the emphasis was on commissioning evaluations based on the demands of donors and country partners. Stakeholders were encouraged to use the draft and feedback on their experiences via an online survey. Uptake was widespread, the draft guidance being used by programme managers, embassy staff, training institutions, evaluation managers, and researchers.

It was employed for major joint evaluations in Sudan (Bennett et al. 2010),

Sri Lanka (Chapman et al. 2009) and the Democratic Republic of the Congo (Brusset et al. 2011). These multi-donor assessments looked at over a decade of support to peacebuilding and statebuilding processes. Norway's evaluation department used the draft to assess the Norwegian contribution to peace in Haiti (Norad Evaluation Department 2009) and Sri Lanka (Norad Evaluation Department 2011). The Swedish military commissioned an evaluation of its Provincial Reconstruction Teams in Afghanistan. At the same time, the European Commission began a review of its peacebuilding portfolio. Germany looked at the impact of aid in Afghanistan (Böhnke et al. 2010) and evaluated its civil peace service programme (Paffenholz et al. 2011). A variety of implementing NGOs used the guidance to evaluate individual projects.

In February 2011 Norway hosted a workshop to review these experiences.[2] This event demonstrated that, while many challenges remain, quality evaluation can provide useful insights to improve donor support to peacebuilding and statebuilding processes.[3]

The body of evidence and policy lessons were widely shared, including at INCAF meetings, to encourage further discussion and debate.[4] Country findings were taken up by donors, implementing partners and recipient governments, and commissioning agencies followed up on specific evaluation findings internally. For example, Canada issued a management response to the multi-donor evaluation in Southern Sudan (OECD 2012). Norway's evaluation of peacebuilding in Sri Lanka was launched with a public debate in Oslo and covered by the Sri Lankan press (Goodhand et al. Forthcoming). The application phase was also critical in building ownership of the final product and supporting development of an active, engaged community of practice working on evaluating peacebuilding and statebuilding interventions. This was seen at the Oslo workshop, which brought together diverse practitioners and policy experts in a constructive, collegial atmosphere.

Throughout the application phase, concerns about the balance between ownership and independence were debated, particularly around involving national and local statekholders — a core tenant of development evaluation. In Sudan the steering committee was co-chaired by the Ministry of Finance of the Government of Southern Sudan and the Evaluation Department of the Netherlands Foreign Ministry. The Congo evaluation involved local communities in conflict analysis. Such experiences usefully demonstrated that participation of partners is feasible, even in the difficult circumstances of violent conflict. But such involvement was problematic elsewhere and sometimes led to accusations of bias. Here the differing views of the peacebuilding and evaluation communities could be seen. For example, when the joint evaluation in Sri Lanka was finalised, peacebuilding experts emphasised its usefulness for in-country discussions, while some evaluators were concerned about the perceived credibility of the evaluation.

At the end of the application phase, it was clear that the guidance had made a number of useful contributions to the field, namely, by bringing greater clarity to key concepts, introducing the use of theories of change as a core pillar for evaluating peace support, describing the use of conflict analysis (in both programming and evaluation), and adapting evaluation criteria to the context of peacebuilding. Users pointed out, however, that the links between these different pieces were not clear. There was also a sense that expectations were too high, in terms of what an individual evaluation could cover and for evaluation producing definitive evidence on 'what works' across incredibly diverse conflict contexts.

These lessons were incorporated into a revised guidance. The final guidance was

less prescriptive, reflecting maturation in the field over time. The document more clearly acknowledges the difficult realities faced by evaluators and at the same time points to the untapped potential of existing research methods. Importantly, it underlines the possibility of maintaining evaluation quality even in conflict settings.

Moving the Agenda Forward

The final guidance was officially launched on 27 November 2012 at a reception hosted by Norway at the OECD in Paris. It provides step-by-step guidance to evaluation and basic principles on good programme design and management.

The guidance should contribute to more and better evaluation and support a more evaluative approach to peacebuilding and statebuilding support — one that encourages critical reflection, strategic thinking and collective learning, and does not hesitate to question untested assumptions about the links between aid, conflict, statebuilding and development. Evaluation is not a substitute for good analysis, solid programme design and strategic management, but should be part of ongoing peacebuilding work, including implementation of the 'New Deal' and agreement on peacebuilding and statebuilding goals with the g7+ fragile states group.

By no means does the guidance solve all of the myriad problems faced when evaluating in settings of conflict and fragility. Data are still often missing and problems of low evaluability and the disconnect between strategic objectives and actual aid allocations seem to remain. Further experience would be useful specifically on using theories of change to evaluate conflict-wide strategies and in thematic (cross-country) evaluations.

The guidance should therefore be seen not as an end itself, but as an input to further experimentation and learning. To support this ongoing development, the DAC

Evaluation Network partnered with Search for Common Ground's Learning Portal for Design, Monitoring & Evaluation for Peacebuilding to create an online discussion (http://www.dmeforpeace.org). Readers are encouraged to participate and to share their own work and experiences.

MEGAN GRACE KENNEDY-CHOUANE is a development professional with expertise in evaluation and peace studies. She works as a policy analyst at the OECD, supporting the DAC EVALNET, and holds a master's degree in public administration and a post-graduate diploma in children, youth and development. The opinions expressed herein are her own and do not necessarily reflect the official views of any organisations with which she is associated.

Endnotes

[1] Challenges have been discussed extensively elsewhere, for example: Smith 2004; OECD & CDA 2007.

[2] Workshop proceedings: http://www.oecd.org/dac/evaluation/dcdndep/workshopone valuatingconflictpreventionandpeacebuilding whathavewelearned.htm

[3] OECD DAC and NORAD (2011) summarise lessons from this phase.

[4] See Kennedy-Chouane 2011.

References

Bennett, J. et al. 2010, *Aiding Peace: A Multi-donor Evaluation of Support to Conflict Prevention and Peacebuilding Activities in Southern Sudan 2005–2010*, Brighton: ITAD.

Böhnke, J. R., **Koeler, J.** & **Zürcher, C.** 2010, 'Assessing the Impact of Development Co-operation in North East Afghanistan 2005–2009: Final Report', Evaluation Report 049, BMZ, Bonn.

Brusset, E. et al. 2011, 'Amani Labda — Peace Maybe: Joint Evaluation of Conflict Prevention and Peace Building in the Democratic Republic of Congo. Synthesis Report', June, Channel Research, Ohain, Belgium.

Chapman, N., **Duncan, D.**, **Timberman, D.** & **Abeygunawardana, K.** 2009, 'Evaluation of Donor-Supported Activities in Conflict-Sensitive Development and Conflict Prevention and Peacebuilding in Sri Lanka — Final Report'. Joint evaluation commissioned by FPS Foreign Affairs, Foreign Trade and Development Cooperation, Belgium; the Ministry of Foreign Affairs, Policy and Operations Evaluation Department, the Netherlands; the Department for International Development (DFID), UK; and the German Federal Ministry for Economic Cooperation and Development (BMZ), Germany.

Goodhand, J., **Klem, B.** & **Sørbø, G.M.** Forthcoming, 'Battlefields of Method: Evaluating Norwegian Peace Efforts in Sri Lanka' in *Evaluation Methodologies for Aid in Conflict*, Abingdon: Routledge.

Kennedy-Chouane, M.G. 2011, 'Improving Conflict Prevention and Peacebuilding Assistance through Evaluation' in *OECD Journal: General Papers* 2010/1, 25 May: 99–107.

Norad Evaluation Department 2009, 'Evaluation of Norwegian Support to Peacebuilding in Haiti 1998–2008', Evaluation Report 5, Norad, Oslo.

Norad Evaluation Department 2011, 'Pawns of Peace: Evaluation of Norwegian Peace Efforts in Sri Lanka, 1997–2009', Evaluation Report 5/2011, Norad, Oslo.

OECD 2012, *Guidance on Evaluating Peacebuilding Activities in Settings of Conflict and Fragility — Improving Learning for Results*, Paris: OECD: http://bit.ly/T4qZrS

OECD & CDA 2007, 'Encouraging Effective Evaluation of Conflict Prevention and Peacebuilding Activities: Towards DAC Guidance', Collaborative Learning Projects, OECD, Paris.

OECD DAC & NORAD 2011, 'Chair's summary — Evaluating Conflict Prevention and Peacebuilding Activities: Lessons from the Application Phase', Workshop, Oslo, 16–17 February: http://www.oecd.org/development/evaluationofdevelopmentprogrammes/dcdndep/48092571.pdf

Paffenholz, T. et al. 2011, 'The German Civil Peace Service: Synthesis Report. Volume I: Main Report', unpublished Evaluation Report, Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung.

Smith, D. 2004, 'Towards a Strategic Framework for Peacebuilding: Getting Their Act Together', Overview Report of the Joint Utstein Study on Peacebuilding, Evaluation Report 1/2004, Ministry of Foreign Affairs, Oslo.

Journal of Peacebuilding & Development

The g7+ and the new deal: an opportunity for south sudanese civil society enhancement

Hafeez A. Wani
Published online: 19 Sep 2013.

PLEASE SCROLL DOWN FOR ARTICLE

The New Deal and the Post - 2015 Development Agenda

THE g7+ AND THE NEW DEAL: AN OPPORTUNITY FOR SOUTH SUDANESE CIVIL SOCIETY ENHANCEMENT

HAFEEZ A. WANI

In the wake of the second anniversary of South Sudan's independence, on 9 July 2013, South Sudanese everywhere must ask the question, 'Is South Sudan heading in the right direction or wrong direction?' This question inevitably inflames debates that reincarnate historic events, commitments and South Sudan's journey towards peace, justice and economic prosperity. It is without question that South Sudan, as the youngest nation in the world, faces very complex and dynamic challenges. At the same time, global support and favour for the new country must challenge South Sudan to make the best use of these opportunities.

This article provides an overview of the New Deal for Engagement in Fragile States (the 'New Deal'), its implementation in South Sudan and the projections of South Sudanese civil society's involvement, contribution and execution of its oversight roles in the process in light of the complex dynamics within the country.

The Emergence of the New Deal

The International Dialogue on Peacebuilding and Statebuilding (IDPS) is a platform made up of the Organisation for Economic Cooperation and Development (OECD) donors, non-traditional donors such as Brazil and China, the United Nations, World Bank, African Union, African Development Bank, representatives of civil society and conflict-affected/fragile states. The IDPS was launched six months after the third High Level Forum meeting (2008) in Accra, Ghana to inform the process of identifying realistic peacebuilding and statebuilding objectives that address the root causes of conflict and fragility at the country level, led by fragile states and supported by development partners.

The IDPS undertook a series of national-level consultations between 2009 and 2010 in Burundi, Central African Republic (CAR), Democratic Republic of the Congo (DRC), South Sudan, Liberia, Sierra Leone and Timor Leste to inform the first International Dialogue meeting in Dili in April 2010. Just prior to the first Dili meeting, a grouping of seven countries affected by conflict and fragility, consisting of Timor Leste, DRC, Cote d'Ivoire, Haiti, Sierra Leone, CAR and Afghanistan, formally established the g7+, an organisation committed to focusing on new ways of engaging to support inclusive country-led and country-owned transitions out of fragility; the group has since expanded to 17, including South Sudan. In Dili, the g7+ issued a statement called 'The Dili Declaration' which challenged donor countries and fragile states to work together to develop an international action plan on peacebuilding and statebuilding — one that

116

would place conflict-affected countries in the driver's seat in diagnosing their own problems, jointly prescribing solutions and carrying out treatment using the peace-building and statebuilding goals (PSGs) as an important foundation to enable progress towards the Millennium Development Goals (MDGs). This was followed by a move to more thematic-based dialogues that in turn formed the substance of the second International Dialogue meeting in Monrovia, Liberia (June 2011), where the five PSGs were agreed upon as preconditions to working towards the MDGs in fragile and conflict-affected states. The product of this meeting was dubbed the 'Monrovia Roadmap'; it highlighted the five PSGs: legitimate politics, security, economic foundations, justice, and revenue and services. This grew into the New Deal, adopted in Busan, South Korea on 30 November 2011 and endorsed by 41 countries and multilateral organisations at the fourth High Level Forum on Aid Effectiveness.

New Deal Implementation in South Sudan

South Sudan is a pilot country for the New Deal (2012–2015), which means that it embarked on the first step of New Deal implementation by undertaking a fragility assessment exercise in August 2012 and holding a structured debate about drivers of conflict, progress and challenges in implementing the PSGs. A draft assessment report was then formulated from the fragility assessment,[1] and a set of indicators were developed to situate South Sudan on the fragility spectrum. In November 2012 a two-day validation workshop was held. The fragility assessment provides an opportunity to present two complementary frameworks, namely a compact for implementing the New Deal and the country aid strategy

South Sudan is one of the most active members of the g7+. The New Deal and membership in the emergent g7+ offer South Sudan an opportunity to address the causes of fragility and build a path towards resilience. Perhaps more importantly, they challenge South Sudan as a country to focus on new ways of engaging with both international and national partners and stakeholders at all levels, including civil society, as well as identifying joint commitments to achieve better results based on common processes and tools of analysis. These commitments are putting many aspects of the country's administrative and political practices to test, with inclusive and participatory political dialogue, relating the local context to the New Deal framework, combating corruption, strengthening the triangular relationship between society, state and donors, and strengthening government structures as donors take a back seat to country-led processes. While progress in many of these areas is already visible, more time, resources, planning and commitment are required to move forward with the pilot implementation of the New Deal.

In August 2012 South Sudan conducted its fragility assessment using a combination of consultations and research aimed at providing a first overview of progress, challenges and priority actions to help the country advance towards resilience. The fragility assessment commenced with a multi-day assessment workshop that brought together 100 participants from 10 state governments, civil society organisations from 10 states, academia and international partners. In focus groups, participants considered drivers of fragility, progress in implementing the PSGs, challenges and priorities. In addition to consultations, the assessment also drew on relevant literature and quantitative data, where available, to illustrate and validate perceptions. The findings and recommendations contained in the first draft were reviewed during a two-day validation workshop in November 2012 attended by more than 50 stakeholders.

117

The fragility assessment workshop was followed by the formulation of a 'Draft Assessment Report' and the first 'Menu of Indicators' to situate South Sudan and assess its progress on the 'fragility spectrum' — a diagnostic tool to support fragile states in the identification of their main weaknesses. To consolidate civil society's efforts in this process a civil society engagement structure was formulated made up of five slots (Civil society Organisations [CSO] country focal point, indicators focal point, implementation focal point, political strategy focal point and reporting focal point) voluntarily occupied by civil society actors. Civil society was represented in the indicators formulation process by the indicators focal point, which later organised a civil society roundtable assessment of the consolidated country indicators in order to better inform the process. Civil society's analysis of the consolidated indicators was well received by the government.

The overall assessment results as articulated in the South Sudan Fragility Assessment draft summary report suggest that the Republic of South Sudan has made sufficient progress on all five PSGs since the Comprehensive Peace Agreement (CPA) interim period and independence on 9 July 2011 to move beyond the crisis stage of the fragility spectrum. While none of the PSGs has yet reached transition, reform efforts seem to have borne most fruit with regard to legitimate politics, following the 99.8% vote for secession from Sudan in July 2011. Moving forward, key challenges include sustainable internal political settlements, the transformation of the security sector, reform of justice institutions, the creation of diversified economic foundations and strengthened capacity for accountable and equitable service delivery.

In April 2013 South Sudan Partnership Conference was held in Washington, DC, organised by the United States government, which, among many things, tabled discussions about fiscal reforms, budgetary support and the New Deal compact formulation, which is an agreement between the government of South Sudan and its development partners (donors) to better consolidate their development efforts with government in the lead. Following this, South Sudan's Ministry of Finance and Economic Planning (MoFEP) led the formulation of a compact steering committee composed of representation from MoFEP, other government institutions like the Ministry of Petroleum and Mining and the Ministry of Cabinet Affairs, donor agencies and civil society. The New Deal compact steering committee is charged with outlining mutually agreed policy benchmarks for the government, matched with commitments from partners to build capacity and improve aid effectiveness, to be completed by September 2013 and signed in conjunction with the South Sudan Investment Conference.

The Opportunity for Civil Society to Grow through Political Engagement

South Sudanese civil society dates back to the 1970s and includes actors from faith-based organisations, advocacy groups, service delivery organisations, relief and emergency organisations, and monitoring groups, among others. The diversity of civil society organisations, their wide geographical coverage and the existence of several thematic network groups are key strengths to be exploited in the implementation of the New Deal. What is needed is an adequate mechanism to allow the civil society to speak with a unified voice that effectively expresses the concerns of citizens from all walks of life. Civil society in South Sudan faces challenges in accessing consistent, predictable institutional funding and maintaining competent human resources — which are often lost to better-paying institutions, therefore leaving a big gap in the organisation's institutional memory.

118

It also faces a disservice in capacity-building initiatives provided by international NGOs, which tend to be project focused and to neglect long-term impact areas like advocacy, resource mobilisation, succession planning, leadership, etc. In spite of these gaps, South Sudanese civil society's strengths lie in its wide geographical distribution, the trust it has built among grassroots communities, and the diversity of sectors of engagement, ranging from highly active sectors like basic service delivery, peacebuilding, human rights and development to less concentrated sectors like environment preservation where its efforts can complement those of government and partners in the implementation of the New Deal.

The New Deal recognises civil society as a principal player in the implementation process and this has provided South Sudanese civil society a platform to support their coordination enhancement initiatives across the ten states with a common focus on national priorities. The fact that the New Deal emphasises the inclusion of civil society in all aspects of analysis and implementation has shifted government–civil-society relations from one of civil society being only a critic of government, to engagement and influencing change as a key player. The New Deal engagement structure is advantageous to civil society on two grounds:

- It has provided a means to begin harmonising civil society's key messages and areas of interest around the five PSGs. With these common objectives, consensus among civil society organisations is more easily generated and efforts towards a common approach to playing an oversight role in the New Deal implementation process are ignited. This has, in the short and medium term, challenged civil society to come up with a common engagement strategy. Ongoing progress includes the formation of a civil society engagement working group, a nationwide concept/plan to strengthen civil society coordination/engagement with the New Deal, establishing an effective north–south civil society relationship, strengthening relations with government (particularly the department of aid coordination in the MoFEP) and securing a civil society seat in the country's compact steering committee.

- It has helped civil society to carve out clear roles it can play in the joint commitment efforts that make up the New Deal implementation process. The debate amidst civil society now is more focused on how to ensure its advisory, monitoring and oversight roles maintain high impact and address the issues of direct concern to citizens. Being a part of a process like the New Deal presents the opportunity for civil society to influence national guidance and reform from within the government systems where the most impact can be made.

The synergy between the g7+, the New Deal and country initiatives such as the South Sudan Development Plan and South Sudan Development Initiative is visible through the goals, objectives and indicators. The development priorities set out in the South Sudan Development Initiative (which is an outcome of the South Sudan Development Plan) target the areas of governance, economic development, social and human development, conflict/security and rule of law. These priorities are the same as the five PSGs of the New Deal, namely, legitimate politics, security, justice, economic foundations, and revenue and services. The non-duplicating nature of the New Deal implementation process further supports harmonisation efforts, as it does not provide room for contradictions in terms of priorities. It adequately serves the objective of creating an enabling environment for economic development in South Sudan.

119

Challenges and Prospects for New Deal Implementation

For a country coming out of conflict with continuing insecurity risks, the success of a process such as the New Deal requires more than just political will; it requires citizen buy-in and support and an increase in government–citizen mutual accountability. Despite the New Deal's emphasis on state–donor accountability, government–citizen accountability is arguably more important in moving a country out of fragility and conflict. Adequacy and the structural necessity of financial management systems, human capital growth and economic diversification especially in the extractive industry and agriculture are vital to development as much as developing a solid local tax base through economic growth. The role of civil society in channelling citizens' voices throughout these avenues is an essential part of the New Deal implementation (through citizen awareness campaigns, harvesting citizen concerns to better inform the government on the impact of their decisions, and citizen auditing of government projects), although it is often not adequately recognised.

In April 2013, Sudan and South Sudan came to an agreement leading to the resumption of oil production, which had been disrupted as a result of a transit fees dispute in January 2012. In the same month 40 countries and international organisations gathered in Washington, DC for the South Sudan partners forum, during which five key agreements were reached: (1) developing a New Deal compact for South Sudan, (2) pursuance of an International Monetary Fund (IMF) staff-monitored programme to strengthen the national macroeconomic framework and reform as well as related budget support from international financial institutions, (3) an EU statebuilding contract to support health and education sectors, (4) a new multidonor partnership fund to strengthen government systems and social service delivery, and (5) support for a private sector investment conference in Juba. These commitments set a good pace for the sustained recovery of South Sudan following the 16-month oil production hiatus. However, only the New Deal compact and private sector investment commitments have civil society involvement; the rest of the commitments leave civil society out, based on the argument that some commitments are particularly relevant to government and its international (donor) partners. This highlights two issues. The first is the continued focus of the government's accountability to international donors with less consideration of accountability to its own citizens. The second issue is that it undermines civil society's oversight and advisory role in the other three commitments. It is essential that civil society is included in these commitments because a well-functioning civil society and politically involved citizenry are the backbone of longer-term sustainable development. Civil society's wide geographical coverage places it in a prime position to play the role of taking the New Deal conversation beyond the state capitals and closer to marginalised groups.

The MoFEP is to be commended for establishing firm budget execution controls for the 2012–2013 fiscal year, including clear procedures for proper authorisations by accounting officers, public expenditure discipline and clear and transparent contracting and payment procedures. A Petroleum Revenue Management Bill is also being debated in parliament, which demonstrates the efforts the government has taken to put in place the necessary structures and policies to support the recovery and growth of the economy. However, the enforcement of regulations and action plans for the budget execution controls and Petroleum Revenue Management Bill do not include strong measures to arraign officials who continue to benefit from embezzling public funds because they think they are entitled to do so as a reward for their participation in the armed struggle leading to independence. The MoFEP has made efforts in consultation with civil society

to find appropriate ways of fashioning civil society's engagement in the public budget process and budget oversight. However, the Petroleum Revenue Management Bill is silent on the role of civil society in monitoring petroleum revenue flow, management of funds and carrying out citizen audits of government petroleum management processes. The level of continued laxity of government in tackling corruption can be construed as threatening to the success of the New Deal, especially in a state where economic and political power are deeply intertwined and the rule of law requires significant improvement. This weakens national ownership and affects the triangular relationship between state, society and donors. Civil society would ideally play the critical role of monitoring budgetary execution, exposing incidences of corruption and advocating for public reparation without fear of being persecuted by state officials.

Similarly in the theme of establishing controls, civil society in South Sudan has begun to refashion its engagement with government through forming specific target-based thematic working groups made up of highly specialised civil society actors to provide objective assessment and analysis of government policies in collaboration with legislators in order to provide a citizen-based perspective on government policies and regulations, for example the civil society working group on the communication bills and the voluntary and humanitarian NGOs bill. This is a practice that has proven useful at the national level and needs to be strengthened at the subnational levels.

Tribal sensitivity is a critical driver of conflict, which was eclipsed by the citizens' focus on a common enemy throughout the civil war. Notwithstanding certain tribal tensions among pastoral communities, it is imperative to reduce the risks of further tribal tensions caused by feelings of unfair distribution of resources, feelings that public offices are dominated by one tribe relative to another, or feelings that there is continued focus on developing basic services in some geographical areas at the expense of others. These tensions and sentiments of dissatisfaction are already present today; to contain the risk of outright conflict, citizen involvement in government development plans must be enforced, so that citizens are aware of the government's efforts to address their concerns and can better understand their role in contributing to economic growth. Fostering an enabling environment for income generation and the growth of a middle class by improving security, infrastructure and communications evenly across the country, and establishing a truth and reconciliation process are areas within which civil society can play a complementary role through awareness campaigns, citizen focus group discussions, harvesting citizens' opinions and presenting them to the government to better inform development processes, and leading by example through adopting friendly intertribal practices like emphasising cross-cultural tolerance in organisational activities, expanding programmes beyond the community of one's descent, etc.

In closing, this analysis lays a foundation for dialogue on four critical questions that lie at the heart of the New Deal project and deserve ongoing attention: Is, and how is, the government of South Sudan in the driver's seat of development in the country? Is civil society role well recognised, integrated and executed? Is citizens' involvement a visible element of the journey South Sudan is undertaking? And is the New Deal the answer to South Sudan's strategy for exit from fragility. The answers to these questions are diverse but what is evident is that the real work of addressing these questions is only just beginning, a redoubling of efforts is needed to ensure that the next steps to implementing the New Deal and progressing out of fragility do not produce lacklustre results like similar processes in the past and this warrants adequate joint efforts from government, civil society and donors.

The burgeoning civil society in South Sudan can lay a foundation to ensure adequate checks will be in place to monitor the progress of the joint commitments and the government's obligation to address the drivers of conflict and fragility. However, the voice of civil society is far from being incontrovertible. This is not to imply that the government of South Sudan does not recognise the importance of civil society or consider its views in its processes. Rather it is to emphasise that the strength of civil society organisations lies in their ability to produce a harmonious force that is legitimate and represents the voices of citizens from all parts of the country through a well-coordinated and targeted strategy.

Endnote

[1] http://www.newdeal4peace.org/wp-content/uploads/2013/05/South-Sudan_FA-Summary_Draft_121212.pdf

An Integrated Vision for Peace and Development? Building on the High Level Panel's Report

Larry Attree
Published online: 19 Sep 2013.

PLEASE SCROLL DOWN FOR ARTICLE

AN INTEGRATED VISION FOR PEACE AND DEVELOPMENT? BUILDING ON THE HIGH LEVEL PANEL'S REPORT

LARRY ATTREE

At the end of May 2013, the High Level Panel of Eminent Persons on the Post-2015 Development Agenda (HLP) published its report 'A New Global Partnership'. The report moved forwards the debate on the global development framework that will replace the Millennium Development Goals (MDGs) when they expire at the end of 2015. Broadly, the report has attracted positive responses, especially from the peace community: it asserts a new development paradigm based on five big, transformative shifts, of which the fourth is the need to 'Build peace and effective, open and accountable institutions for all':

> We must acknowledge a principal lesson of the MDGs: that peace and access to justice are not only fundamental human aspirations but cornerstones of sustainable development.

This has set the stage for a global conversation at the highest level about the relationship between peace and development, which is long overdue. But, beneath the headlines, is the HLP report good for peacebuilding? This briefing summarises three positive points about the HLP report, flags three areas requiring further thought, and notes some of the challenges looking forward.

How Has the HLP Done?

Almost all key peacebuilding issues are included

Comparing the goals and targets from the HLP report with the targets for addressing conflict and violence Saferworld set out in February (see the table annexed to Saferworld's latest briefing paper [Saferworld 2013, 11–12]), we are fairly satisfied that, notwithstanding one or two gaps, the HLP has done a good job of including most critical global priorities in its illustrative goal-and-target framework — in a way that corresponds with the global evidence of the key challenges as well as what works in addressing them.

Addressing inequalities

As part of its agenda to ensure that 'no one is left behind', the HLP calls for targets upheld by indicators that are disaggregated in new ways:

> Data must also enable us to reach the neediest, and find out whether they are receiving essential services. This means that data gathered will need to be disaggregated by gender, geography, income, disability, and other categories, to make sure that no group is being left behind.

123

This responds positively to the call by the UN Peacebuilding Support Office and Saferworld (Brinkman et al. 2013), and a range of other peacebuilding orgaisations (Civil Society Platform for Peacebuilding and Statebuilding 2013), to ensure the framework focuses the minds of policymakers on inequalities between social groups in all areas of public life — in order to reduce the potential for such inequalities to feed into the grievances that in turn lead to conflict and violence in so many contexts around the world.

Peace is integrated into the HLP development paradigm as a dimension

The HLP's report and illustrative goals and targets framework is also impressive in that it grasps the argument for integrating peace as a dimension: i.e. including a goal that focuses on overcoming aspects of violence, insecurity, and injustice that do not fit into other thematic areas, as well as integrating other building blocks of sustainable peace across the new framework. This proposal for integrating peace across the new framework as a 'transformative shift' could initiate 'a decisive move towards coherence between actors and sectors and between local, national and global solutions' that 'could bring important multiplier effects for both development and peacebuilding effectiveness' (Saferworld 2013, 5).

Areas for Development

Although there is a clear need to build on the ideas of the HLP to accomplish a new policy consensus that integrates peace into the vision for sustainable human development for the next generation, there are inevitably areas that will require further discussion.

Are the interconnections across the framework clear enough?

One question for those hoping for a framework that helps to measure whether countries are making progress towards sustainable rather than negative peace concerns interconnections between peace-related issues across the new framework. Designers of the post-2015 accountability framework need to take care to ensure that measurement of progress towards reduced violence and greater security (under HLP Goal 11) should not become separated from measurement of progress in other areas that are crucial to sustained peace and contained under other goals (in particular HLP Goal 10 on governance).

Informal justice, dialogue, and social cohesion are missing?

Peace, security, and justice are clearly *in* focus, but this focus is on formal institutions whereas informal justice, dialogue, and dispute resolution are left out. These should be part of the post-2015 framework if it is to be comprehensive. A suggested target in this area, 'Divisions within society are constructively resolved', and illustrative indicators that could be further developed to uphold it, are set out by Saferworld elsewhere (2013, 11).

Are counter-terrorism and stabilisation up for debate?

Similarly, when it comes to the 'external stresses that lead to conflict', alongside the range of issues that the HLP has raised in this area (illicit flows of drugs, arms, precious minerals, money, etc.), a truly open conversation on this topic should include discussion

of the security agendas of developed countries. There is a need to critically examine current approaches to counter-terrorism and stabilisation, and talk through alternative, less militaristic, and more developmental approaches to solving these issues.

Where Next? Political and Technical Sides of the Conversation

The HLP report will now need to be digested within a number of multilateral processes leading, presumably, towards an agreement by UN member states on a post-2015 sustainable development framework in 2015. The HLP's work will feed into a range of discussions that will play a role in shaping the post-2015 framework: the upcoming Special Event towards achieving the Millennium Development Goals in September 2013; the deliberations of the Open Working Group on Sustainable Development Goals, which discusses conflict in February 2014 before reporting to the General Assembly later in the year; the high-level political forum on sustainable development which has its first meeting in September 2013; and the work of various expert groups working on technical aspects of the new framework, such as indicators and financing. A report by the Secretary General is expected to draw all these strands together in 2014. The diplomatic negotiations around the post-2015 framework, and the policy dialogue that grows around them, will bring about some intense contestation between different visions and paradigms for future development. This will at times be difficult, and may not lead to the acceptance that peace is both a requirement for development progress and an end in itself. Nonetheless, it will be an opportunity for states to challenge one another's views on how the world's most pressing problems can best be solved. As all states have significant room for improvement in their approaches to crisis-ridden and violence-affected contexts, the conversation by itself could prove incredibly valuable to engendering a more conflict-sensitive international community, in an increasingly multipolar world.

While member states have the opportunity to form a common view of how to do better in reaching development and peacebuilding outcomes, this will need to be linked to the technical conversation on the accountability framework (i.e. the elaboration of measurable indicators to underpin the post-2015 goals and targets). The direction of travel set out by the HLP is broadly progressive, but many of its targets are vague. Therefore the difficult work must now begin to define more clearly targets and indicators for goals 10 and 11, as well as other key targets relevant to peacebuilding, with credible indicators that are technically and politically feasible.

Based on an identification of over 160 existing multinational metrics of relevant issues, Saferworld argued in February 2013 that, although sustainable peace is a multidimensional concept, measuring progress towards addressing the key challenges that lead to conflict and violence is nonetheless feasible if political consensus is reached around an accountability framework and capacities are developed in the right areas.

The technical conversation around indicators is already underway, the UN convening groups of relevant experts to discuss options. Existing analyses discussing indicators on peace, security, justice, and governance issues tend to agree on certain key departure points regarding peacebuilding indicators.[1] The most important of these is that no single indicator can in every context tell a full, fair story about progress. In the post-2015 framework, a clear danger to be avoided is of targets and indicators being established that distort priorities by measuring inputs as opposed to outcomes. As argued by Saferworld, drawing on a number of other analyses, peacebuilding indicators can only work well if

JOURNAL OF PEACEBUILDING & DEVELOPMENT

they measure three key aspects of a target: *capacity* to address the issue at stake; the *'objective' situation* in society; and, crucially, the *perceptions* of all social groups on security, justice, and other key peace-related issues. Of these, measuring public perceptions of progress is the most crucial.

This thinking can be applied to the HLP report. The HLP's illustrative targets 11b and 11d are as follows:

> 11b. Ensure justice institutions are accessible, independent, well-resourced and respect due-process rights;
> 11d. Enhance the capacity, professionalism and accountability of the security forces, police and judiciary. (HLP 2013, 31)

While it is very positive that these targets on security and justice themes have been included, at the same time these targets are more focused on strengthening capacities and less on achieving whole-of-sector outcomes than they should be — in contrast with the avowed focus on outcomes in the HLP's illustrative framework. In the move from the HLP report towards an accountability framework, security and justice targets need to be clearly defined in terms of achieving (human) security and justice for all social groups, and indicators need to be agreed for these targets which include a focus on people's perceptions of whether they are secure and whether justice is being done.

Alongside the technical work to be done on developing indicators that offer a practical and reliable way to measure progress towards peace, at strategic level there needs to be swift financial and political investment in a 'data revolution' to support the transformative global development framework proposed by the HLP. Capacity for measuring progress which is global, regular, confidential, impartial, and trusted needs to be developed. At present, many issues relevant to peace are monitored by Western research and advocacy organisations. The indicators they currently measure would perhaps be more acceptable to all UN member states if they were taken up within multilateral monitoring and statistical bodies. Indeed, the technical conversation about post-2015 indicators needs to be a bottom-up process that learns from the existing efforts by the African Union and the g7+ to pilot the monitoring of peace, security, justice, and governance indicators: political consensus is not impossible if the experience and ideas of all member states are taken into account in a genuinely open global conversation about how peace relates to development.

Endnote

[1] Cf. Scheye 2009, 7, 13, 16–17, 19; UN 2010, 40; United Nations Development Programme (UNDP) 2007, 12. On the importance of public confidence and perceptions surveys see also World Bank (2011).

References

Brinkman, H.-J., **Attree, L.** & **Hezir, S.** 2013, 'Addressing Horizontal Inequalities as Drivers of Conflict in the Post-2015 Development Agenda', UN PBSO, Saferworld, Columbia, February: www.saferworld.org.uk/downloads/Inequalities-conflict-FV.pdf

Civil Society Platform for Peacebuilding and Statebuilding 2013, 'Joint Submission to Consultation on Shared Indicators', January: www.saferworld.org.uk/downloads/CSO-joint-submission-to-consultation-on-shared-indicators.PDF

HLP 2013, 'A New Global Partnership', UN.

Downloaded by [University of York] at 23:14 26 September 2013

Saferworld 2013a, 'Addressing Conflict and Violence from 2015: A Vision of Goals, TARGETS and Indicators', www.saferworld.org.uk/downloads/Post-2015-4th-goals,targets-and-indicators-FINAL.pdf

Saferworld. 2013, 'Addressing Conflict and Violence from 2015: From the High Level Panel to an Accountability Framework', June: www.saferworld.org.uk/downloads/pubdocs/From-the-High-Level-Panel-report-to-an-accountability-framework.pdf. Accessed August 2013.

Scheye, E. & **Chigas, D.** 2009, 'Development of a Basket of Conflict, Security and Justice Indicators', May: http://www.scribd.com/doc/23391272/Conflict-Secuirty-Justice-Global-Indicatiors-Final-Report

UN 2010, 'Monitoring Peace Consolidation — United Nations Practitioners' Guide to Benchmarking': http://www.un.org/en/peacebuilding/pbso/pdf/monitoring_peace_consolidation.pdf. Accessed August 2013.

UNDP 2007, 'Governance Indicators: A Users' Guide', 2nd Edition: http://www.undp.org/content/dam/aplaws/publication/en/publications/democratic-governance/oslo-governance-center/governance-assessments/governance-indicators-2nd-edition/governace_indicator_undp_users_guide_online_version.pdf. Accessed August 2013.

World Bank 2011, *World Development Report 2011*, Washington, DC: World Bank.

Journal of Peacebuilding & Development

Resources

Published online: 19 Sep 2013.

PLEASE SCROLL DOWN FOR ARTICLE

RESOURCES

Book Notices

Conflict Assessment and Peacebuilding Planning: Toward a Participatory Approach to Human Security

Lisa Schirch

Lynne Reiner 2012

ISBN 978-1-56549-578-4

Offering a systematic approach that links practical conflict-assessment exercises to the design, planning, monitoring, and evaluation of peacebuilding efforts, *Conflict Assessment and Peacebuilding Planning* has been carefully — and realistically — designed to enhance the effectiveness of peacebuilding practice. A logical framework and tools for improving planning and implementation are presented, while the analysis provides a basis for questioning widely held assumptions about conflict assessment and peacebuilding planning. This concise handbook, informed by on-the-ground realities, is an essential resource for building true human security.

Context-Sensitive Development: How International NGOs Operate in Myanmar

Anthony Ware

Kumarian Press 2012

ISBN 978-1-56549-523-4

Focusing on Myanmar, with its perfect storm of extreme poverty, international sanctions, and egregious political repression, Anthony Ware shows how context sensitivity can help development organisations to better meet the needs of their client populations. Ware points out that, while practitioners have questioned universal economic prescriptions for development, they have been less rigorous in questioning the normative foundations behind their work. Though Myanmar is his case in point, he suggests key issues of perception and practice that are intrinsic to a successful development enterprise wherever it is undertaken.

Peacebuilding through Community-Based NGOs: Paradoxes and Possibilities

Max Stephenson and Laura Zanotti

Kumarian Press 2012

ISBN 978-1-56549-427-5

Max Stephenson and Laura Zanotti explore the contested, but increasingly relevant, role that nongovernmental organisations (NGOs) play in resolving conflict and bringing about peace and security in the global arena. The authors draw on case studies from Haiti, Serbia, and Northern Ireland to highlight the range of ways that NGOs are involved in postconflict reconstruction efforts. In the process, they not only explore the outcomes and effects of various past strategies, but also caution strongly against one-size-fits-all approaches to peacebuilding and offer food for thought about the complexities facing NGOs and international donors as they engage in postconflict situations.

Time to Listen: Hearing People on the Receiving End of International Aid

Mary Anderson, Dayna Brown, and Isabella Jean

Collaborative Learning Projects 2012

ISBN 0988254417

Time to Listen represents the cumulative evidence of five years gathering evidence from people living in societies that are recipients of international aid. CDA's Listening Project organised teams of 'listeners' across 20 countries and contexts to gather the voices, insights, and lessons from people both inside and outside the aid system. This publication represents the lessons that have come forth through conversations with nearly 6,000 people. Using their words, their experiences, and their ideas, it describes why the cumulative impacts of aid have not met expectations and describes a way forward to make changes that, according to those on

130

the receiving end, will lead to more effective results.

Reports

Peacebuilding with Impact: Defining Theories of Change

CARE International UK 2012

http://www.international-alert.org/
sites/default/files/publications/
120123CAREDefiningTheoriesChange_
FINAL.pdf

This 2012 report argues for the benefits of theories of change in improving the effectiveness of peacebuilding interventions. More attention to theory of change can clarify project purpose and intended results, identify the stakeholders, make the gaps between local and national more apparent, highlight ineffectiveness, enhance conflict sensitivity, make the need for conflict analysis clear, and aid more efficient collaboration. This paper is a summary of the findings of research that assessed 19 peacebuilding projects in three conflict-affected countries: Uganda, Nepal, and the Democratic Republic?of the Congo. Individual country-specific papers are also available and a 'How to Guide to Using Theories of Change in Peacebuilding' is forthcoming.

OECD DAC Fragile States 2013: Resource Flows and Trends in a Shifting World

Organisation for Economic Cooperation and Development's Development Assistance Committee 2013

http://www.oecd.org/dac/incaf/
resourceflowstofragilestates.htm

This report focuses on fragile states, the social, economic and political causes of state fragility, and their local national and global origin. Drawing on ten central questions about the trends, theory, and future trajectory of fragility, the report divides into three chapters. The first describes the evolution of fragility as a concept, the second analyses financial flows within states, and the third covers issues that are likely to be the drivers of fragility in the future. The report concludes by highlighting the 2011 New Deal for Engagement in Fragile States, which called for greater consideration among international partners of the particular conditions that fragile states face while incorporating peacebuilding and statebuilding goals into their respective intervention programmes. In its conclusion, the report calls for future research to deepen and consider both internal and external stress factors when evaluating state fragility.

A Peacebuilding Tool for a Conflict-Sensitive Approach to Development: A Pilot Initiative in Nepal

Asian Development Bank 2012

http://www.adb.org/publications/pea
cebuilding-tool-conflict-sensitive-approa
ch-development-pilot-initiative-nepal

Socioeconomic and political inequalities were central to the decade-long conflict in Nepal. Now, Nepal's transition from a fragile state to a nation capable of human development has brought a new initiative from the Asian Development Bank to help build sustainable peace. These initiatives focus on postconflict situations with respect to which the report utilises analytical tools to benefit project leaders in assessing the issues, implementing development projects, and processes for mitigation. The strategy is built around utilising the peacebuilding tool to analyse risks and prepare recommendations.

Rural Women in Peace-Building and Development in Nigeria

Ambily Etekpe

The Dawn Journal 2012

http://www.thedawnjournal.com/wp-
content/uploads/2011/07/8-Ambily-
Etekpe.pdf

In 2005, the Ekpetiama Clans and the Nigeria Liquefied Natural Gas company worked together to create an integrated oil and gas project that would benefit industry and the surrounding Bayelsa State in the southern Niger delta region. The integrated oil and gas project was a multi-billion-dollar proposal that was

supposed to bring about 47 different programmes that would benefit community development. The Shell Petroleum Development company failed to deliver on those promises and in 2010 this caused a revolt started by the rural women's community. This report by Ambily Etekpe focuses on David Easton's theory of 'post behavioural revolution', in which the 'participation observation' method is utilised to explore the need for rural women to advocate their knowledge of peacebuilding and development in a male-dominated vocation. This 2012 report in *The Dawn Journal* discusses the lessons learned from the revolt, the need for dialogue, the importance of the Global Memorandum of Understanding, and the responses of government and the Shell Petroleum Development company to prevent escalation. The report recommends the continuation of a micro credit scheme that has positive impacts on rural communities.

Building Just Societies: Reconciliation in Transitional Settings, Workshop Report

Enrique Sánchez and Sylvia Rognvik

United Nations 2012

http://www.un.org/en/peacebuilding/pbso/pdf/12-58492_feb13.pdf

Reconciliation is a necessary step in the movement towards sustainable peace and development. Society must build healthy relationships between citizens, between different groups in society, and between the state and society. This report derives from a workshop organised by the UN Peacebuilding Support Office, the Norwegian Peacebuilding Resource Centre, and the Kofi Annan International Peacekeeping Training Centre. It highlights important thematic areas of reconciliation such as reparation, healing, and truth and justice. It also examines countries engaged in reconciliation process, and the international community's involvement, and offers reflections and suggestions on the future of reconciliation.

Dialogue Series 10: From Peacebuilding and Human Development Coalitions to Peace Infrastructure in Colombia

Borja Paladini Adell

Berghof Foundation 2012

http://www.berghof-handbook.net/documents/publications/dialogue10_paladini_adell_comm.pdf

The Berghof Handbook Dialogue Series chooses case studies relevant to conflict transformation, integrating both scholarly analysis and practitioner experience. The tenth edition deals with peace infrastructure in Colombia based on the experience of a United Nations Development Programme (UNDP) practitioner. The paper discusses the conflict situation in Colombia and the need for local peace infrastructure. It looks at the limits of the national peace infrastructure and the need to begin initiatives from below. At the local level in Colombia there exists great potential for peacebuilding and development with the actors, capacities, and constituencies available. However, the real challenge lies in how to combine and utilise these resources for sustained and productive peacebuilding efforts. The author argues for a form of hybridity in which local peace infrastructure is primary, complemented by international actors within a larger framework with a combination of short-term and long-term goals for peacebuilding and sustainable human development.

Impact Evaluation Conflict Prevention and Peacebuilding Interventions

Marie Gaarder and Jeannie Annan

World Bank 2013

http://econ.worldbank.org/external/default/main?pagePK = 64165259&piPK = 64165421&theSitePK = 469372&menuPK = 64166093&entityID = 000158349_20130624113015

The international community is paying increased attention to the 25% of the world's population that lives in fragile and conflict-affected settings, acknowledging that these settings represent daunting development challenges. To deliver better results on the

ground, it is necessary to improve the understanding of the impacts and effectiveness of development interventions in contexts of conflict and fragility. This paper argues that it is both possible and important to carry out impact evaluations even in settings of violent conflict, and it presents some examples from a collection of impact evaluations of conflict prevention and peacebuilding interventions. The paper examines the practices of impact evaluators in the peacebuilding sector to see how they address evaluation design, data collection, and conflict analysis. Finally, it argues that such evaluations are crucial for testing assumptions about how development interventions affect change — the so-called 'theory of change' — which is important for understanding the results on the ground.

The Future of Intrastate Conflict in Africa: More Violence or Greater Peace?

Jackie Cilliers and Julia Schünemann

Institute for Security Studies 2013

http://www.issafrica.org/futures/uploads/Paper246Edit.pdf

This paper analyses future trends for intrastate conflict in Africa up to 2050 using the International Futures model. After reviewing the main post-Cold-War patterns of conflict and instability on the continent, the paper discusses seven key correlations associated with intrastate conflict in Africa. It then points to a number of reasons for the changing outlook, including the continued salience of various 'structural' conditions that drive intrastate violence even during rapid economic growth, and recent improvements in human development alongside a strengthened regional and international conflict prevention, conflict resolution, and peacebuilding regime. Finally, the paper explores how multipolarity may impact on stability and forecasts trends for intrastate conflict in West, Southern, East, and Central Africa. The authors expect large-scale violence to continue its steady decline, although the risk of instability and violence is likely to persist, and even increase in some instances.

E-communications

DM&E for Peacebuilding

http://www.dmeforpeace.org/.

Created for Design, Monitoring and Evaluation (DM&E) professionals, *DM&E for Peacebuilding* (www.dmeforpeace.org) is a website that provides evaluation reports and data, information about best and emerging practices, methodologies, and new events and opportunities. The website brings together new information in the field while also providing a number of forums for peacebuilding practitioners to discuss general issues and network. *DM&E for Peacebuilding*'s Resources section provides a number of reports and documents for professionals. Arranged by theme, subject, and type, the resources provided range from case studies and evaluation reports to opinion pieces and lectures. Each resource is given a brief description and a PDF of the resource is provided. The Bookmark section allows users to set aside relevant resources for reading and use later on. Users can upload documents, start a discussion with other users, and post available opportunities. The website's Opportunities section provides links to job opportunities, events, a roster of consultants, and specific funding opportunities. Registration is free and required for access to information. A User section lists those who have joined the site and allows registered users to reach out to one another. Registration is open to both established professionals and those looking to enter the field.

Documents

Somalia Conference 2013: Final Communiqué

https://www.gov.uk/government/news/somalia-conference-2013-communique

The Conference discussed issues related to the political situation, security, justice and policing, public financial management, funding, stabilisation, refugees and internally displaced persons, and the role of multilateral organisations and international support. The participants

recognised that this was only one in a series of ongoing conferences about the future of Somalia.

Excerpt from the communiqué:

> The Somalia Conference took place at Lancaster House on 7 May 2013, co-hosted by the UK and Somalia, and attended by fifty-four friends and partners of Somalia.
>
> We met at a pivotal moment for Somalia. Last year Somalia's eight-year transition ended and Somalia chose a new, more legitimate Parliament, President and Government. Security is improving, as Somali and AMISOM [African Union Mission in Somalia] forces, and their Ethiopian allies, recover towns and routes from Al Shabaab. The number of pirate attacks committed off the coast of Somalia has drastically reduced. The famine has receded. The diaspora have begun to return. The economy is starting to revive.
>
> But many challenges remain. Al Shabaab is still a threat to peace and security. The constitution is not complete. Piracy and terrorism remain threats. Millions still live in Internally Displaced Persons and refugee camps. The country lacks developed government structures, schools, hospitals, sanitation and other basic services.
>
> The Federal Government of Somalia has set out its plans to address these challenges in its Six Pillar Policy. At the Conference, the international community came together to agree practical measures to support the Federal Government's plans in three key areas — security, justice and public financial management. The Federal Government presented its vision for the implementation

of federalism, the adoption of a permanent constitution and holding of elections. We also agreed to work together to tackle sexual violence in Somalia.

> We agreed that partnership between Somalia and the international community would form the basis of our future cooperation: the international community is committed to provide coordinated and sustained support for implementation of the Federal Government's plans.
>
> … The Conference agreed that Somalia had made significant progress. We congratulated all who had made that possible, notably the Somali people, Federal Government, Members of Parliament, civil society and diaspora. We commended the sustained commitment of Somalia's international partners, and urged continued results-orientated support. We recognised the need to consolidate progress quickly and reiterated our determination to support Somalia over the long-term.

The Horn of Africa region's growing strategic and economic importance has made establishing a more effective central government a global priority. As Somalia progresses, the country's relationship with the outside world will become even more important. The conference was a significant step towards strengthening the relationship between Somalis, the UK, and other key international partners.

Key Conclusions of the Global Thematic Consultation on Violence, Citizen Security and the Post-2015 Development Agenda

http://uncsd.iisd.org/news/panama-consultation-discusses-violence-security-in-post-2015-agenda/

From 31 January to 1 February 2013, over 100 UN officials, civil society representatives, and stakeholders met in Panama City for a consultation on 'Violence,

Citizen Security, and the Post-2015 Development Agenda'. The meeting considered the linkage of violence to citizen insecurity, HIV, disasters, and conflict and how these linkages affect the ability of countries to achieve the 2015 Millennium Development Goals (MDGs) and to further sustainable development. Emphasis was placed on conflict, crime, and violence and their implications for the MDGs in Latin America and the Caribbean.

The final document that emerged from the meeting included key considerations that acknowledge that, while progress has been made towards a number of the 2015 MDGs, violence threatens the stability of these gains. The prevention and reduction of violence has the potential to advance the gains seen in areas such as the empowerment of women, primary education enrolment, and the reduction of child mortality, and countries are now calling for issues of security and justice to be part of the post-2015 MDG agenda.

The document calls for the next generation of global development goals to 'reduce violence, and promote freedom from fear and sustainable peace'. Proposed targets on the way to meeting these goals include the prevention and reduction of violent deaths, the prevention and reduction of violence among vulnerable groups, and addressing deeper causes of violence such as lack of education and unemployment. In these efforts, the document noted that capacity should be built for information collection and monitoring based on the lessons learned in Latin America and the Caribbean.

Finally, participants called for 'the reduction of violence and promotion of freedom from fear and sustainable peace' to be made a key pillar of the post-2015 development agenda.

135