Routledge
Taylor & Francis Group

Journal of Peacebuilding & Development

ISSN: 1542-3166 (Print) 2165-7440 (Online) Journal homepage: http://www.tandfonline.com/loi/rjpd20

Evaluation in Conflict Zones: Methodological and Ethical Challenges

Kenneth Bush & Colleen Duggan

To cite this article: Kenneth Bush & Colleen Duggan (2013) Evaluation in Conflict Zones: Methodological and Ethical Challenges, Journal of Peacebuilding & Development, 8:2, 5-25, DOI: 10.1080/15423166.2013.812891

To link to this article: http://dx.doi.org/10.1080/15423166.2013.812891

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) 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?

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

*We would like to thank the following people for their contributions to refining our thinking about the issues addressed in this paper: Colin Knox; Rick Davies; the students of the evaluation modules we taught at the INCORE Summer School (2010–2012); the participants in workshops on ethics in conflict zones in the Community of Evaluators Meeting (Kathmandu, February 2013) and the African Evaluation Association Meeting (Accra, January 2012); and two anonymous peer reviewers.

JOURNAL OF PEACEBUILDING & DEVELOPMENT, VOL. 8 No. 2, 2013

ISSN 1542-3166 PRINT/2165-7440 ONLINE
http://dx.doi.org/10.1080/15423166.2013.812891

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

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

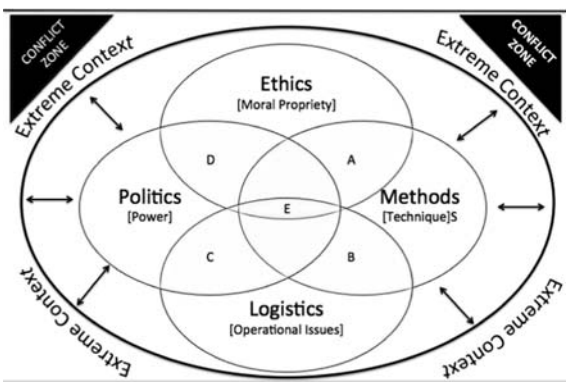
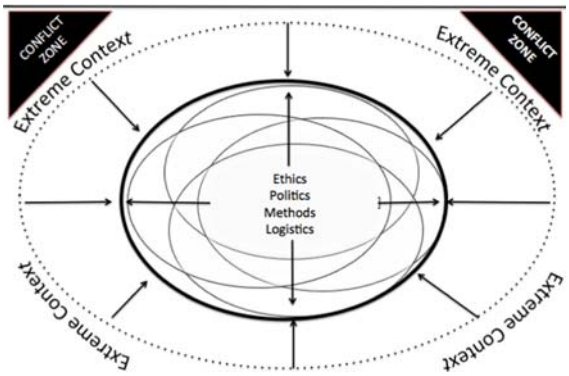


Figure 2: The Amalgamation of Evaluation Domains in Conflict Zones as Intensity Increases



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.

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?'

> '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.

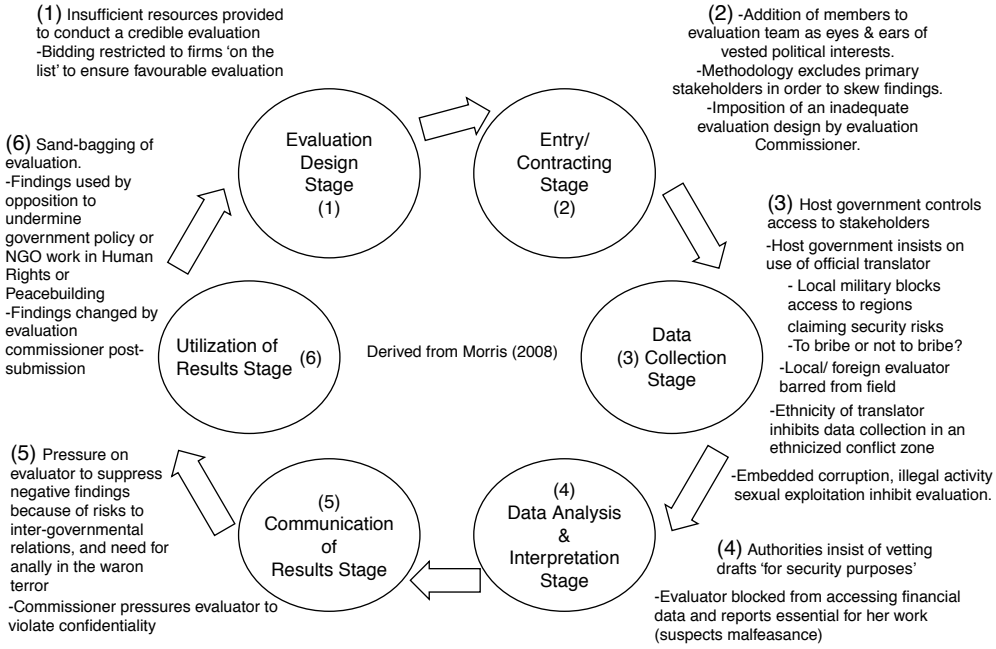
Ethics, unlike methodology, do not easily lend themselves to standardisation; no two situations are identical, and we are dealing with human morality and human behaviour.

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



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.

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

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/dialogue1_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-handbook.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.