BEYOND MITIGATING RISKS AND DISRUPTIONS: EXPANDING THE MEANING OF CONFLICT SENSITIVE EVALUATION

Promising Practices and Practical Considerations for Designing and Conducting Conflict Sensitive Evaluations

March 2017, updated February 2019

Developed by Isabella Jean, Henry Nelson, and Elisabeth Ris

The Peacebuilding Evaluation Consortium (PEC)

The Peacebuilding Evaluation Consortium (PEC) is a project of Alliance for Peacebuilding (AfP) in partnership with CDA Collaborative Learning Projects, Mercy Corps and Search for Common Ground (SFCG). The project is funded by the Carnegie Corporation of New York (CCNY) and is field-wide effort to address the unique challenges to measuring and learning from peacebuilding programs. The PEC convenes donors, scholars, policymakers, local and international practitioners, and evaluation experts in an unprecedented open dialogue, exchange, and joint learning. It seeks to address the root causes of weak evaluation practices and disincentives for better learning by fostering field-wide change through three strategic and reinforcing initiatives: 1) Developing Methodological Rigor; 2) Improving the Culture of Evaluation and Shared Learning; and 3) Fostering the Use of Evidence to Inform Peacebuilding Policy.

CDA Collaborative Learning Projects

CDA (www.cdacollaborative.org) is a US based non-profit organization committed to improving the effectiveness of those who work internationally to provide humanitarian assistance, engage in peace practice, support sustainable development, and conduct corporate operations in a socially responsible manner. CDA helps peace practitioners and organizations improve the relevance and accountability of programming through better tools for conflict analysis, program strategy, design, and monitoring and evaluation. In the evaluation sphere, CDA has contributed to influential policy guidance, such as the OECD/DAC guidance on evaluating conflict prevention and peacebuilding activities.

Acknowledgements

Table of Contents

Section I: Introduction

A range of approaches, tools, and tactics exist to mitigate risks to implementers and evaluators evaluating interventions in fragile and conflict-affected settings. At the same time, evaluations are interventions in and of themselves and can affect dynamics in the contexts they examine just as they can be affected by such dynamics. In fragile contexts, there is a heightened risk that evaluations, as interventions, can exacerbate conflict drivers inadvertently, causing more violence. Given this potential, there is a "need to redress the relative absence of the systematic consideration and incorporation of conflict context into the theory and practice of evaluation."[1]

Recognizing evaluations as interventions, members of the Peacebuilding Evaluation Consortium have expressed concern that evaluation choices have implications for inter-group relations and trust among the parties involved as evaluands. Additionally, PEC members note that evaluators are rarely asked to identify and address ways in which their presence and methodological choices may affect conflict dynamics and that there is little agreement on how to do so efficiently and effectively.

This paper was developed in response to that identified need for more resources on conflict sensitivity. The paper examines practices used to limit potential negative effects on evaluation processes in fragile and conflict-affected contexts and explores emerging practices for understanding and reducing the extent to which evaluation processes and products negatively impact local and national dynamics in divided fragile contexts.

Section II: What is Conflict Sensitive Evaluation? Why does it matter?

To be conflict sensitive in an evaluation process[2] is to be deliberate about recognizing and minimizing potential the risks and harmful effects of the evaluation on group and conflict dynamics and mitigating risks posed by the conflict to beneficiaries, evaluation participants, and the evaluation team. However, practitioners across the peacebuilding and development fields have not always conceptualized conflict sensitive evaluation this comprehensively. Over the past two decades, there has been limited agreement on what makes an evaluation "conflict sensitive."

While there is a general understanding in the field of peacebuilding and conflict resolution that evaluations of programs *should be* conflict sensitive, the phrase "conflict sensitive [monitoring and] evaluation" is not used uniformly by all evaluators, practitioners and policymakers. For example, some evaluations that reference the need for conflict sensitive evaluation refer to a need for M&E that specifically supports peacebuilding activities and or M&E of interventions delivered in conflict areas[3] or

---

[1] Duggan, Colleen, and Kenneth Bush. "Lessons for Researchers and Evaluators Working in the Extreme." In *Evaluation in the Extreme: Research, Impact and Politics in Violently Divided Societies*, edited by Kenneth Bush and Colleen Duggan, Thousand Oaks: SAGE Publications, 2015, p.6.

[2] Throughout this piece, we emphasize the entire process of conceptualizing, designing, conducting, reporting on, and disseminating evaluation findings as each phase has the potential to both be affected by conflict and to affect conflict dynamics.

[3] For example, ACT for Peace—a UNDP peacebuilding program in Mindanao, Philippines—released a report calling for more "conflict sensitive M&E" for peacebuilding initiatives. However, this definition of conflict-sensitive M&E is less about the ways that conflict can affect evaluations and vice versa, but rather is used interchangeably with "peace-promoting M&E." In this context, conflict-sensitive M&E refers to M&E that specifically supports peacebuilding and conflict resolution efforts. For more, see Villanueva, Starjoan D. "Managing Performance in Peacebuilding: Framework for Conflict-Sensitive Monitoring and Evaluation." *GoP-UN Act for Peace Programme* (2009).

M&E that considers how dynamics change over time and the interactions between the project and those dynamics (but, critically, not the interaction between the M&E activities and conflict dynamics).[4]

Those interpretations of "conflict-sensitive M&E" share key attributes, including the need to build a careful understanding of the conflict context into the M&E design and processes. As discussed above, the evaluation and implementation communities are increasingly aware of a need to take conflict sensitivity a step farther by considering how M&E designs and processes affect the conflicts and dynamics that they seek to examine.

In response to the call to expand our collective conceptualization of conflict sensitive M&E, our review of relevant resources and practical tips from evaluators focuses on conflict sensitive M&E as a means of: 1) understanding the conflict context in which the evaluation is taking place; 2) understanding how the evaluation process (all stages) might unintentionally exacerbate conflict dynamics and how evaluators can minimize these effects; and 3) understanding the challenges the conflict contexts present to the evaluation and how evaluators effectively address these challenges. The following three sections explain the rationale for in-depth consideration of each factor.

Understanding the conflict context in which the evaluation is taking place

In the available literature on conflict sensitive evaluations, there is relative agreement about the importance of an adequate conflict analysis. In order to properly understand the context before designing and conducting the evaluation, evaluators must systematically analyze the conflict. An accurate and nuanced conflict analysis allows "a detailed, politically informed, anthropological understanding of the social, cultural, economic and political structures and processes within the project environment—in addition to the usual set of technical evaluation skills expected of a professional evaluator."[5]

Incorporating a conflict analysis into an evaluation is essential for two reasons. First, the conflict analysis allows the evaluator to adequately assess how well the program adapted to the conflict context. The original conflict analysis forms the basis for all the assumptions on which the program is predicated, making it imperative for the evaluator to understand the conflict analysis that was conducted.[6]

Second, the conflict analysis helps the evaluator adapt the evaluation to the conflict context in order to ensure that the evaluation is conflict sensitive.[7] Third, the conflict analysis is a foundation on which the evaluation designers can begin to build an understanding of how the evaluation could affect group, inter-group, and conflict dynamics.

[4] For example, a Saferworld module on conflict sensitivity differentiates conflict-sensitive M&E from "traditional monitoring and evaluation" by noting that the former requires "an understanding of the context as it changes over time [and] measuring of the interaction between the project and the context." Saferworld. "Chapter 3, Module 3: Conflict-sensitive monitoring and evaluation." For more, see In *Conflict-sensitive approaches to development, humanitarian assistance and peace building: tools for peace and conflict impact assessment* (2004), p.1.

[5] Duggan and Bush, p. 303

[6] Broegaard, Eva, Beate Bull and Jens Kovsted. "Evaluation approaches in situations of conflict and fragility." In Evaluation Methodologies for Aid in Conflict, edited by Ole Winckler Andersen, Beate Bull, and Megan Kennedy-Chouane, 15-37. Oxford: Routledge, 2014., p. 25

[7] Barnett, Chris and Jon Bennett. "Critical reflections on the South Sudan evaluation of conflict prevention and peacebuilding activities." In Evaluation Methodologies for Aid in Conflict, edited by Ole Winckler Andersen, Beate Bull, and Megan Kennedy-Chouane, 38–60. Oxford: Routledge, 2014, p.48

How the evaluation process affects conflict dynamics

The process of carrying out targeted, in-depth research and analysis, and using that information to generate conclusions and recommendations for ongoing and future interventions (i.e. carrying out an evaluation) can have positive and negative consequences on the social, military, political, and economic dynamics at play in a conflict. The following sections focus on three domains[8] of evaluation practice (political, ethical, and logistical/methodological) that can significantly affect the extent to which an evaluation negatively impacts a conflict environment.

Political Domain

Even when an evaluation or research project is methodologically sound, "we need to understand the political and societal contexts within which research is embedded and through which impacts are mediated. We need to ask, therefore, how individuals and institutions appropriate (or misappropriate) research, and apply (or misapply) it for the purposes of influencing policy or practice."[9] Particular ways that evaluation can effect conflict in the political realm include the risk of evaluators becoming "pawns" of political forces in complex conflict environments[10] and the risk of raising expectations of communities participating in the evaluation.

Evaluators can also increase the chances that their efforts will drive conflict by sending implicit political and ethical messages. For example, the presence or the actions of evaluators might provide false hope or raise the expectations of communities participating in the evaluation. Communities might expect more aid or programming due to implicit messages that evaluators impart, or due to the mere presence of the international evaluator in their community. According to Goodhand, by promoting unrealistic expectations, evaluators can inadvertently drive conflict dynamics.[11]

In addition to inadvertently raising expectations, evaluators can implicitly "[give] the impression that they are legitimizing warring groups and analyze who may or may not be gaining political capital out of their activities."[12] For example, if an evaluator negotiates access from a rebel group or government forces occupying a combat zone, are they implicitly conferring legitimacy on these groups? How can evaluators balance the need to stay safe with the desire to remain neutral in conflict zones? Compounding this issue is the so-called "Hawthorne effect"— "the evaluation may also exert an influence on the political dynamics and, as a result, influence the conflict by involving contending forces in ways that enhance their legitimacy."[13]

---

[8] In this and subsequent sections we borrow from the "Domains of Evaluation" framework devised by Bush and Duggan. While this framework focuses on the ways in which the conflict context can affect evaluations, it can also be a useful tool for determining how evaluation can affect conflict dynamics. Bush and Duggan consider the effects of conflict across four specific domains of evaluation: (1) evaluation logistics; (2) evaluation methodology; (3) the politics of evaluation; (4) ethical concerns in evaluations. This document summarizes potential effects of evaluations on conflict across these domains (combining domains (1) and (2) into a single logistics/methodology domain). It should be noted that these three domains are by no means exclusive: they overlap and influence each other in important ways.

[9] Bush and Duggan book, 5

[10] Goodhand, Jonathan, Bart Klem, and Gunnar M. Sorbo. "Battlefields of method: evaluating Norwegian peace efforts in Sri Lanka." In Evaluation Methodologies for Aid in Conflict, edited by Ole Winckler Andersen, Beate Bull, and Megan Kennedy-Chouane, 61-84. Oxford: Routledge, 2014, p. 61

[11] Ibid, 14

[12] Ibid.

[13] OECD. "Encouraging Effective Evaluation of Conflict Prevention and Peacebuilding Activities: Toward DAC Guidance." OECD Journal of Development 8:3 (2007): 7-97, p. 29

There is an inherently political nature to peace processes and peacebuilding that can permeate the evaluation stage of programs. Politics can interfere with the evaluation in two broad ways, which the OECD refers to as "politics with a little p" and "politics with a big P."[14] Politics with a little p refer to political decisions made regarding the evaluation itself, including the composition of the evaluation team, the methodology used and who is included in data collection. Politics with a capital P refers to actual political decisions that are generally made by the donor or implementing agency.

Political interference in evaluations in fragile and conflict-affected states can occur at any stage in the evaluation process, including evaluation commissioners pre-determining what an evaluation will conclude;[15] commissioners trying to influence or control evaluation findings and recommendations; suppression of evaluation reports; and changing the TOR or other evaluation methodological documents – to limit access to those who may have unfavorable views of the intervention or otherwise skew results - mid-evaluation without agreement from the evaluator.[16] According to one survey of evaluators, the stage of the evaluation where evaluators were most likely to face political interference was during the communication of results, with commissioners and evaluands pressuring evaluators to portray a program in an untruthful (usually more positive) way.[17]

How evaluators cope with political interference can have serious effects on conflict dynamics. For example, a commissioner or evaluand may attempt to influence or bury evaluation reports that find that illegal, unethical, or dangerous acts were committed during program activities.[18] In fragile and conflict-affected states, these acts could include program activities that contribute to conflict. Therefore, a "conflict-sensitive evaluation" might also be one that accurately reports on the failure or lack of conflict sensitivity in the program itself. This raises the question: in order to be conflict sensitive, do evaluators have an obligation to report on conflict sensitivity shortcomings of the program they are evaluating? According to Morris, an evaluator must follow up on his or her work to ensure that the evaluation is used correctly.[19] Can evaluations only be conflict sensitive if they identify and provide recommendations to remedy conflict-driving aspects of the evaluand? And furthermore, to what extent is the evaluator required to ensure that these results are used? Few evaluators outside of government accountability offices have leverage vis-a-vis the commissioning organization to ensure that evaluation findings are used correctly, and such practices are rarely documented.

Ethical Domain

Ethical considerations are relevant across the evaluation process and must be examined throughout. In reflecting on the rollout of Save the Children UK's revised evaluation guidance and the procedures put in place to support ethical considerations, Benelli and Low note that "Doing ethical research is not something that can be secured once and for all: it is a continuous process, as well as an interaction between the researcher and the research subjects."[20]

---

[14] Ibid, 48

[15] Morris, Michael. Evaluation Ethics for Best Practice: Cases and Commentaries. New York: Guilford, 2007.

[16] Turner, David. "Evaluation Ethics and Quality: Results of a Survey of Australasian Evaluation Society Members." AES Ethics Committee (2003): 1-5.

[17] Morris.

[18] Ibid

[19] Ibid

[20] Benelli, Prisca and Tamara Low. "The Ethics of Doing Primary Research with Vulnerable Populations: Facing the Challenges One Step at a Time." Paper submitted to Forced Migration Review, February 2019.

According to Jayawickrama and Strecker, "... it is imperative to start [an evaluation] with what one knows and recognize one's inherent values and bias."[21] This process can be understood as a form of "cultural humility," which is the assumption that all of an individual's life experiences and affiliations add up to a complex way of viewing the world. Adopting cultural humility is different from cultural competency, because the former requires the evaluator to undergo "on-going self-evaluation and self-critique" in order to understand one's own values.[22] According to Jayawickrama and Strecker, while cultural humility is important for all evaluators, it is especially important for evaluators in violently divided societies. Causing an evaluation participant to deviate from his or her cultural norms in the process of the evaluation can put the participant at risk in fragile or conflict-affected states.[23]

Vulnerable groups are defined by Jayawickrama and Strecker as either individuals who are being asked sensitive questions about sexual behavior, illegal or political history, experience with violence, or gender/ethnic status; or individuals who can only be reached through a gatekeeper, such as a parent or a community leader.[24] While vulnerable groups are important beneficiaries of many peacebuilding programs, "evaluation frameworks provide little guidance for how to approach vulnerability."[25] Due to this lack of guidance, there is a real chance that evaluators expose vulnerable people to greater risk or exacerbate existing inequalities.

Connected to issues of vulnerability is the risk that an evaluation might retraumatize individuals by posing sensitive questions. Evaluators are challenged to obtain objective information while simultaneously implementing a "lite-touch" approach to data-collecting that does not retraumatize the vulnerable.

Logistical and Methodological Domain

An evaluation's data collection processes can impact conflict dynamics and expose participants to risks. Even when community members have given their informed consent, their answers to sensitive questions might put them at risk. Similarly, questions that are appropriate and even benign in one context might endanger or anger respondents in another context.[26] Evaluators can inadvertently expose individuals to risk by identifying someone as a "gatekeeper" or "community leader" while conducting an evaluation. A community member's closeness to the evaluation can prove dangerous if parties to the conflict see that person as a potential rival or threat.[27] Similarly, if the evaluator identifies one individual or group as community leaders and uses these individuals as gatekeepers to reach the broader community, they might risk alienating and angering rival leaders in the community.[28]

---

[21] Jayawickrama, Janaka, and Jacqueline Strecker. "The Ethics of Evaluating Research: Views from the Field." In Evaluation in the Extreme: Research, Impact and Politics in Violently Divided Societies, edited by Kenneth Bush and Colleen Duggan, 127-164. Thousand Oaks: SAGE Publications, 2015, p. 133
[22] Ibid, 134
[23] Ibid, 138
[24] Ibid, 139
[25] Ibid.
[26] Saferworld, p.5
[27] Goodhand, 12
[28] Ibid

Challenges that conflict contexts present to the evaluation

Logistical and Methodological Domain

This section will outline two broad logistical and methodological dilemmas that evaluators face when making evaluations conflict sensitive in fragile and conflict-affected states: (1) issues pertaining to access; (2) issues pertaining to security and conflict.

According to Ford, lack of infrastructure and ongoing violence can greatly limit the ability to conduct research by limiting evaluators' access.[29] Limits to access might continue even after conflict has ceased—for example, a team of evaluators researching the effects of Norway's peace mediations in Sri Lanka was refused visas by the Sri Lankan government.[30] Because of the inability to access certain areas or populations, evaluations in conflict zones can be "biased, incomplete, and/or (voluntarily or involuntarily) censored," resulting in assessments that do not adequately evaluate the program.[31]

Confidentiality, anonymity and informed consent are generally considered essential for ethically conducting research on human subjects. In the "Three Ethical Principles" for research on human subjects, these concepts are enshrined in "Respect for Persons" principle: the need to acknowledge the autonomy of research subjects and protect research subjects from diminished autonomy.[32] Upholding this principle is essential for protecting evaluation participants and ensuring that the evaluation doesn't do harm. While ensuring respect for individuals participating in the evaluation is essential across all evaluation contexts, failing to ensure confidentiality and anonymity in conflict environments can be deadly.

While protecting those who participate in an evaluation should be a top priority of evaluators, there is also the need to protect those who conduct and assist in the evaluation. Failing to build an appropriate security protocol for the evaluation team and interlocutors can have serious negative impacts on conflict dynamics. Indeed, Mazurana and Gale note that "researchers who lack the knowledge, ability, and/or discipline to make good decisions to stay physically and emotionally healthy and safe are a risk to themselves, the other people on their team, and the people they are interviewing."[33]

Section III: Promising Practices for Conflict Sensitive Evaluation

This section describes a range of methods and approaches currently being deployed to enable conflict sensitive evaluation as well as the critiques and challenges associated with them. The following methods and approaches are described: conflict analysis; considerations for evaluation design, data collection, and adaptation; technology-enabled M&E processes; Third Party Monitoring; and Institutional/Ethics Review Boards. In analyzing these practices, this section draws from both academic and

---

[29] Ford, Nathan, Edward J Mills, Rony Zachariah, and Ross Upshur. "Ethics of conducting research in conflict settings." Conflict and Health 3:7 (2009): 1-9.

[30] Goodhand, Klem and Sorbo, 71

[31] OECD. "Evaluating Peacebuilding Activities in Settings of Conflict and Fragility." OECD Publishing: DAC Guidelines and Reference Series (2012): 7-60. p, 32

[32] Campbell, Susanna P. "Literature Review: Ethics of Research in Conflict and Post-Conflict Environments." Program on States and Security, The Graduate Center, City University of New York (2010): 1-11. p, 1

[33] Mazurana, Dyan, and Lacey Andrews Gale. "Preparing for Research in Active Conflict Zones: Practical Considerations for Personal Safety." In Research Methods in Conflict Settings, edited by Dyan Mazurana, Karen Jacobsen, and Lacey Andrews Gale, 279-291. Cambridge: Cambridge University Press, 2014. p, 277

donor/implementer guidance as well as conflict sensitive practices gleaned from evaluations of programs in fragile and conflicted-affected contexts.

Conflict Analysis Processes and Practical Steps

Along with the need to conduct a conflict analysis, there are a variety of other methods suggested by evaluators and research for making evaluations conflict sensitive. These methods can roughly be broken down into four categories: pre-evaluation logistics; evaluation design; data collection and analysis; and dissemination and follow-up.

Pre-evaluation logistics

The composition of the evaluation team is a key starting point as working with a diverse team will help the evaluation avoid bias, whether real or perceived.[34] It widely considered to be good practice to have a mix of international and local evaluators on an evaluation team, and peacebuilding programs require a further level of specification in choosing evaluation team members. To reduce real and perceived bias, evaluators and enumerators should be chosen to reflect the various ethnic, religious, and linguistic groups associated with the conflict.[35] Bias can skew the results of the evaluation by ignoring the intervention and evaluation's effects on specific communities. Both bias and perceived bias can lead to worsening conflict dynamics if communities feel marginalized or ignored by the results of the evaluation.

While putting together a diverse and representative evaluation team seems relatively straightforward, this process can be complicated in practice. For example, in an evaluation of a conflict resolution intervention in Sri Lanka, a number of national evaluators had to recuse themselves for political reasons. Ultimately, the evaluation team chose not to include *any* Sri Lankan researchers because the selective withdrawal of some members threw off the ethnic balance of the team. According to evaluation team leaders, "this suggests that the principles of independence and transparency, considered to be essential pillars of 'good practice' in evaluation, may be very difficult to abide by in highly politicized and conflict environments."[36]

The timing of the evaluation is another key consideration. Evaluation activities should be timed to avoid intersecting with politicized dates and flashpoint events. For example, evaluators should examine the likely consequences of conducting interviews during an election where violence between rival political parties is expected. While the Saferworld module on conflict sensitivity notes that tensions can make acquiring accurate information difficult,[37] conducting an evaluation under this condition could also drive conflict if evaluators' questions increase hostility between groups that are already experiencing enmity due to a flashpoint event. Avoiding potentially destabilizing events means that evaluators need to have either intimate knowledge of the context or work closely with the program and other local actors who do. OECD guidance spells out the following questions for the evaluator, evaluation manager, and commissioner to consider during the planning stage:

- What is happening in the situation of conflict and fragility? At what stage is the conflict cycle? Watch carefully for potential conflict triggers (elections, controversial celebrations, etc.)

---

[34] OECD Evaluating Peacebuilding, 54.
[35] Goodhand.
[36] Goodhand, Klem and Sorbo, 71
[37] Saferworld, p. 16.

- Would an evaluation at this moment be disruptive to the policy, project, or program itself?

- Would an evaluation spark a political reaction that could undermine the intervention by calling attention to it or by inadvertently feeding political forces in opposition to it?

- Would an evaluation put stakeholders at personal or political risk? Will there be sufficient access to stakeholders, or sufficient safeguards, to avoid bias that might endanger the policy, program or project and the staff and stakeholders?[38]

Evaluation design

One way that evaluation designs can mitigate potentially contributing to conflict drivers is by selecting research participants from across communities in conflict. For example, a review of a dozen relevant evaluations identified a promising strategy for reducing the chance that the evaluation would lead to intra-community violence. In a 2013 impact evaluation of a gender-based violence reduction project in Cote d'Ivoire, the target group was only married women. However, because the intervention included an economic empowerment component that would benefit participants in the impact evaluation, the evaluators chose to include non-married women as well and to choose from this larger pool using a lottery system. While the evaluation simply notes that this methodology was chosen to "preserve social cohesion,"[39] another piece by one of the evaluators of the Cote d'Ivoire project elaborated that methodology was specifically chosen to reduce the risk of intra-community violence. Because the villages where the evaluation was taking place had recently experienced ethnic violence, the evaluators reasoned that picking a smaller pool of women in a non-lottery fashion might prompt suspicion and even conflict.[40]

According to some evaluators, adaptation and flexibility are important aspects of a conflict sensitive evaluation design because they allow evaluators to cope with unforeseen events that could increase the risk that the evaluation will cause conflict. For example, the evaluation team might need to adapt the sample or the survey timing if increasing tensions in an area make data gathering untenable.[41] When rationales are well documented and implications for the usability of evaluation findings are made explicit, adaptability and flexibility is generally well received by donors and implementing agencies who are cognizant of their own ethical and safeguarding policies.

Data collection and analysis

During the data collection and analysis stage of an evaluation, there are several methods for making evaluations more conflict sensitive. First, it is important to obtain truly informed consent from participants. If interviewees do not understand the purpose of the research or the evaluation, they might not fully understand that they are at risk of violence when the results of the evaluation are released. The first step for ensuring informed consent is to build a relationship predicated on mutual respect and honesty between the evaluator and the interviewee. Building this relationship requires the evaluator to honestly examine the power relations between the evaluator and members of the

---

[38] OECD Evaluating Peacebuilding, 50

[39] Gupta, Jhumka, et al. "Gender norms and economic empowerment intervention to reduce intimate partner violence against women in rural Cote d'Ivoire: a randomized controlled pilot study." BMC International Health and Human Rights 13:46 (2013).

[40] Gaarder, Marie, and Jennie Annan. "Impact evaluation for peacebuilding: challenging preconceptions." In Evaluation methodologies for aid in conflict, edited by Ole Winckler Andersen, Beate bull, Megan Kennedy-Chouane, Routledge, 2014. pp.130-154, 139.

[41] Ibid, 136.

communities in which he or she is gathering data.[42] This includes "communicating respectfully and openly with participants and community members throughout the project, respecting autonomy and life-style, and providing useful debriefing about the nature, findings, and value of the research and its likely dissemination."[43] Additionally, in order to obtain informed consent, the evaluator should make clear the risks and benefits of participating in the evaluation, along with explaining the exact role of the evaluator and the evaluation.[44] This can be difficult because the level of risk that the interviewee will be exposed to is likely unknown. Calculating risk is easier and more accurate when the evaluator has access to an up-to-date and accurate conflict analysis.[45]

There are some interesting examples from the field of academic research of innovative methods for ensuring informed consent in conflict zones. For example, while conducting research among illiterate, conflict-affected people in El Salvador, Wood used an oral (as opposed to written) consent protocol. This protocol allowed participants who were illiterate or semi-literate to participate in and understand the purpose of the research. Additionally, the oral consent protocol allowed the researcher to bypass keeping a written record of participants, which could jeopardize their well-being.[46]

After consent has been secured, evaluators must ensure that their data collection methods are conflict sensitive. For example, interview questions must be carefully worded in order to avoid "[shaping] people's perceptions of the conflict," which could in turn "reinforce distrust and hostility toward the 'other side.'"[47]

Additionally, evaluators and enumerators must ensure that they ask questions in a manner that does not increase conflict and tension by harming vulnerable populations.[48] This is often an ongoing process, as evaluators are responsible for ensuring that all members of the evaluation team and enumerators consistently uphold good practices articulated in data collection protocols. Negative cases abound: Jayawickrama relates a case from Darfur, where researchers at an IDP camp badgered their research subjects into talking about their past suffering, rather than their current conditions. To Jayawickrama, "such stories illustrate the myopia of researchers who selectively seek and instrumentally use, information that suits preconceived notions while ignoring realities, problems and needs of the community within which they are working."[49] In another incident in Darfur, an evaluator in a refugee camp asked a young woman's mother to leave the room so the evaluator could speak to the young woman about her rape at the hands of a militia. The mother was extremely angry and immediately ended the interview.[50]

These examples are illustrative of some of the key lessons and good practices for evaluation in peacebuilding environments. First, the evaluator's humility and respect for participants is essential. Evaluators must recognize their own biases and the tendency of some evaluators to "selectively seek and instrumentally use, information that suits preconceived notions while ignoring realities, problems

---

[42] Goodhand, 14.
[43] Campbell, 2.
[44] Conflict Sensitivity Consortium. "How to Guide to Conflict Sensitivity." Conflict Sensitivity Consortium (2012): 4-46.
[45] Ibid, 3.
[46] Wood, Elisabeth. "The Ethical Challenges of Field Research in Conflict Zones." Qualitative Sociology 29 (2006): 373-386. p, 380
[47] OECD Evaluating Peacebuilding, 35
[48] Broegaard, Eva, Beate Bull and Jens Kovsted. "Evaluation approaches in situations of conflict and fragility." In Evaluation Methodologies for Aid in Conflict, edited by Ole Winckler Andersen, Beate Bull, and Megan Kennedy-Chouane, 15-37. Oxford: Routledge, 2014. p, 25
[49] Jayawickrama and Strecker, 147
[50] Ibid, 142

and needs of the community within which they are working."[51] Second, an understanding of the local culture is essential for conflict-sensitive evaluations. Researchers and evaluators must have a "nuanced understanding of local conditions, and an awareness of whom to talk to, how to speak to them, and on what topics."[52] With this understanding, evaluators can interpret the principles of Do No Harm through a local lens to properly understand ways in which the evaluation might drive conflict.

As noted by Benelli and Low, avoiding duplication of research is a key means to limit re-traumatization and respect research participants' time, privacy, and personhood. Conducting secondary reviews and sharing data – with adequate protections in place – are two ways to reduce the burden on research participants. These processes are not as effective as they could be. "With regards to data utilisation, poor knowledge management and high turn-over of humanitarian personnel can limit the full potential of data use."[53]

Dissemination and follow-up
Finally, the aforementioned examples demonstrate the tension between the two tasks of the evaluator in fragile or conflict-affected context: to report accurate information back to the evaluation commissioner and users, and to ensure that the evaluation does not harm participants and communities. According to Jayawickrama, "evaluators must therefore balance their obligation to report the truth, while balancing their responsibility to stakeholders and participants to prevent harm..." [54]

Overcoming that tension largely falls on the personal and professional integrity of the evaluation team members. To support evaluators' commitments to integrity, several of the national/regional evaluation societies have articulated guiding principles that evaluators can draw upon to explain their positions and advocate for good practices in the face of disagreement between evaluators and evaluands or commissioners.   For example, the American Evaluation Association principle on integrity states that "Evaluators behave with honesty and transparency in order to ensure the integrity of the evaluation" and specifically states that evaluators: "Accurately and transparently represent evaluation procedures, data, and findings" and "Clearly communicate, justify, and address concerns related to procedures or activities that are likely to produce misleading evaluative information or conclusions."[55]

Technology-Enabled M&E Processes
Increasingly, in conflict-affected and traditional development contexts, technology is being used to improve the efficiency and reliability of M&E. There is a need to better understand the implications of increased use of Information and Communication Technology for M&E (ICT4M&E) on the conflict sensitivity of evaluation.

For example, the use of technology for monitoring and data gathering in Somalia has "steadily increased" in recent years, with the most common iterations of the trend being handheld devices that allow for quicker collection and analysis of data. Unfortunately, "security, privacy and confidentiality [are] not being adequately addressed" in ICT4M&E.[56] Because there are no established ethical guidelines for using technology in M&E, it can be difficult for evaluators to assess how well they are upholding ethical principles. Additionally, in some contexts, the use of ICT can put evaluators, data collectors and

---

[51] Ibid, 147

[52] Campbell, 4

[53] Prisca Benelli and Tamara Low. Paper submitted to Forced Migration Review. 2019.

[54] Jayawickrama, 149.

[55] American Evaluation Association. 2018 Updated Guiding Principles.

[56] Raftree, Linda, and Michael Bamberger. "Emerging Opportunities: Monitoring and Evaluation in a Tech-Enabled World." The Rockefeller Foundation Evaluation Office (2014): 1-46, p. 40.

participants at greater risk than paper forms. Individuals using phones, tablets and other devices, might be more likely to be accused of being journalists or spies, which in a "highly militarized and politically charged environment" could be very dangerous for evaluators and those with whom they are working.[57]

In contrast, in contexts where smart phone usage is common practice, tech-enabled tools can be invaluable, provided that when comprehensive training precedes the use of the tools. For example, recent programming in Syria saw that enumerators could submit data that had been collected electronically and then delete it from their devices before crossing through an area with active checkpoints, improving their safety. In addition, devices can be programmed to rapidly delete or hide data in case of seizure. In addition, where dynamics shift rapidly, electronic data is faster and cheaper to process – meaning that analysis and programmatic adaptation can happen quicker, which could improve an intervention's ability to address conflict.

Of note, while organizations are increasingly sharing M&E data with each other, some scholars have noted that the ethical dilemmas associated with open data have not been adequately explored.[58]

Third Party Monitoring

Remote management has become a common tool for donors, INGOs, and other major development actors hoping to mitigate the risk that programming poses to both staff and beneficiaries.[59] In parallel, remote monitoring, enabled by third party research entities, has become increasing popular. According to the World Bank, Third Party Monitoring (TPM) is "monitoring by parties that are external to the project or program's direct beneficiary chain or management structure to assess whether intended outputs, outcomes, and impacts have been achieved by the project."[60] While TPM is often used to ensure an unbiased and impartial measurement of delivery and impact of a project, it is increasingly used in fragile and conflict affected settings when program staff and evaluators are prevented from accessing certain insecure areas.[61]

In a review of accountability and learning across programming in Somalia and Afghanistan, the Securing Access in Volatile Environments (SAVE) program found that TPM programming has several advantages for programs implemented in insecure settings. In Somalia and Afghanistan, it allowed donors to (1) increase monitoring (and accountability) in insecure areas; (2) allowed low-visibility monitoring to reduce risk of violence in areas with limited access; (3) allowed more frequent monitoring than would be possible with staff alone; and (4) provided independent triangulation of data to ensure accountability from partners, allowing donors and implementers to verify outputs.[62] These advantages have the potential to both improve the precision and efficacy of M&E, and make M&E more conflict sensitive by reducing the tensions and risk inherent in program staff and evaluators entering insecure areas to conduct research.

However, the SAVE analysis found that TPM had several drawbacks that could reduce the quality of M&E and indeed make M&E processes *less* conflict sensitive. TPM can have a negative impact on conflict sensitive M&E in two ways: (1) TPM providers might lack knowledge and skills in conflict sensitive and ethical data collection techniques; and (2) organizations contracting TPM providers often

---

[57] Ibid, 40-41.

[58] Jayawickrama, 142.

[59] Schultz, Christina. "Remote management is key to keeping aid workers safe in conflict areas." Devex, August 24, 2015.

[60] World Bank, How to Notes: Participatory and Third-Party Monitoring in World Bank–Financed Projects: What Can Non-State Actors Do? DEALING WITH GOVERNANCE AND CORRUPTION RISKS IN PROJECT LENDING.

[61] Ober, Heidi. "The What and Why of Third-Party Monitoring." The Guardian, May 6, 2015.

[62] SAVE Accountability and Learning, 5

do not sufficiently consider the ethical concerns related to transferring risk to TPMs. "TPM providers had little awareness of humanitarian principles and their behavior in the field sometimes created serious reputation risks for the commissioning agencies."[63] The SAVE report largely frames this as a "reputational risk" for the donor contracting the TPM.[64] However, such a narrow definition of this risk ignores some of the larger implications, including the possibility that TPMs, just as local implementing partners might post risk and inadvertently cause harm. This potential negative consequence is compounded by the fact that many TPMs are based in the regions in which they conduct monitoring. While their linkages to certain areas allows TPMs more access than other monitors or evaluators, it can also make them seem less impartial or neutral, which can lead to mistrust or tensions. In order to address this problem, some donors are increasing their conflict sensitivity training for TPMs and are starting to take into consideration the perception of local contractors and other organizations. An additional consideration is the concern that in hiring individuals living in areas heavily affected by violent conflict, TPM's may incentivize them to stay in that area, exposing them to heightened risks.

In addition to the lack of conflict sensitivity skills and knowledge on the part of the TPMs, the contracting agencies themselves often fail to embrace conflict sensitivity when hiring outside monitors. While donors and implementing organizations often hire TPMs to reduce the risk to their staff, the SAVE analysis found that "only one contracting agency consulted has taken this transfer of risks into account in its own procedures and assumed responsibility for the security of monitoring missions."[65] By failing to adequately address the risks the level of risk transferred to the TPMs, contracting agencies are failing to consider the dangers posed to the TPMs and the communities in which they are gathering data.

To sum up, while TPM can make M&E more effective and conflict sensitive in a number of ways, the lack of consistency in ethical practice across TPMs and the failure of many contracting agencies to adequately account for the risks posed by using TPMs make their use risky from a conflict sensitivity perspective. The successful use of TPMs in M&E should be further explored to determine best practices for reducing risk.

Using IRBs and Other Ethics Review Boards

Western research institutions began instituting ethics review boards in the 1980's. Medecins Sans Frontieres (MSF) led the way its field, instituting an IRB process in 2002. A decade later, USAID put in place requirements that research involving human subjects be reviewed by an ethics review committee and DFID and ECHO have instituted research policies that cover the topic of ethics. NGOs are now beginning to catch up.[66] Over the past five years, other organizations (including Marie Stopes, Save the Children, and the International Rescue Committee) have begun using IRBs and ERBs to assess the ethicality of evaluation designs.

MSF uses an internal ethics review board with a unique framework for assessing the ethics of a research or evaluation design for a conflict setting. While more geared toward research than evaluations, the MSF framework is built around three broad benchmarks:[67]

Benchmark 1: Collaborative Partnership

---

[63] Ibid, 6

[64] Ibid, 36

[65] Ibid, 38

[66] Benelli, Prisca and Tamara Low. Paper submitted to Forced Migration Review, February 2019.

[67] Ford, Nathan, Edward J Mills, Rony Zachariah, and Ross Upshur. "Ethics of conducting research in conflict settings." Conflict and Health 3:7 (2009): 1-9. p, 7

- Are the researchers collaborating with other national and international researchers as much as is relevant and appropriate?
- This benchmark might be difficult to achieve when all available partners are overly partisan or in other ways not idea collaborators.

Benchmark 2: Community Engagement

- Do researchers respect the community's values, culture, traditions, and social practices? Are community members involved in the design and implementation of the research?

Benchmark 3: Social Value

- Is the social value of the research increased by ensuring that knowledge is disseminated and that the community benefits from that knowledge?

Some evaluators find that they are better able to cope with ethical challenges when they have access to IRBs. In a survey of members of the Australasian Evaluation Society, Turner found that evaluators will at times draw from more stringent university guidelines (i.e., IRB guidelines) when conducting evaluations where ethical concerns are palpable.

Jayawickrama and Strecker agree that IRBs can often be useful by "[affording] researchers an opportunity to think through ethical scenarios that could manifest and develop mitigation strategies to reduce risk."[68] However, they also address three key drawbacks of using IRBs:

1. Sometimes, and especially in fragile or conflict-affected states, the ethical considerations of the IRBs are not stringent *enough* to prevent possible harm.
2. Researchers and evaluators often use ethical standards from their home institutions, which might not be relevant or appropriate in the fragile or conflict-affected context.
3. The academics who serve on IRBs often have no experience working in fragile and conflict-affected countries.
4. Instead of relying on IRBs and ethical standards from the United States and Western Europe, Jayawickrama and Strecker advocate for mixing the ethical standards and professional codes of the evaluation field with the local norms of the evaluation context.[69]

---

[68] Jayawickrama and Strecker, 128
[69] Ibid, 151

Section IV: Questions for Further Inquiry

1. What are the requirements and buildings blocks for an enabling environment for conflict sensitive evaluations? What is needed to secure funding and political buy-in on the institutional/donor-level to support implementers and evaluators in carrying out conflict sensitive evaluations?

2. Many of the methodological concerns and techniques identified for implementing a conflict sensitive evaluation are focused on the basics: ensuring that evaluation participants have provided informed consent, providing them with anonymity, and keeping sensitive evaluation data confidential. While each of these techniques are certainly important and contribute to a conflict sensitive evaluation, are there other, more complex methodological best practices that evaluators should be taking advantage of (for example, the method used in the Gupta 2014 evaluation)?

3. In general, existing guidance on conflict sensitive evaluation focuses on the negative side of conflict sensitivity: how can the evaluator avoid doing harm. However, there is another side to evaluation as well—a successful conflict sensitivity effort both reduces the intervention's exacerbation of conflict dynamics (dividers) while also strengthening the forces that promote cohesion and peace (connectors). Few authors focused on the second half of the conflict sensitivity equation. Goodhand mentions the need to "do some good" as a researcher, and notes that research can help counter myths and stereotypes and overturn elites' domination of the information economy. According to Goodhand, breaking down barriers and reducing stereotypes and distrust can improve future responses to conflict in communities.70 How might this work for evaluations? How much should the evaluator focus on strengthening connectors and local capacities for peace in the course of his or her evaluation?

4. What additional modifications to the traditional IRB model are needed to make IRBs useful tool for developing conflict sensitive evaluations?

5. The resources we reviewed noted that the traditional conflict analysis and conflict sensitivity tools are not particularly useful for evaluators. Do evaluators agree? Are there better tools that exist for evaluators? If not, what might a more useful tool look like?

6. Are there most recent promising practices for conflict sensitivity emerging with the use of new tools and methods in the evaluation field, including TPM and ICT4M&E?

7. To what extent do evaluation managers and commissioners split the task of creating a conflict sensitive evaluation with the evaluators themselves? How much should the evaluation scope and TOR require in the way of conflict sensitivity and how much should the evaluator be expected to develop?

---

70 Goodhand, 14.

Section V. Questions to Guide Considerations for Conflict Sensitive M&E Processes

Commissioning an Evaluation
(Note, many of the questions relevant to commissioning an evaluation also affect the processes of developing analysis and conclusions, as well as reporting on and disseminating findings.)

Commissioners:	▪ How might potential findings and conclusions from this evaluation affect the lives of the people participating in either the evaluation and or intervention as well as those living in communities affected by the intervention and the evaluation?
	▪ What processes are in place to respond to findings of deliberate wrongdoing or inadvertent negative programmatic consequences?
	▪ What processes are in place to protect evaluators who identify credible allegations of wrongdoing?
Commissioners and Implementers:	▪ Thinking ahead to the end of the evaluation process, what processes are in place to communicate evaluation findings to research and program participants in a manner that supports shared understanding and mutual trust between communities and implementers and amongst communities themselves?
	▪ Would an evaluation at this moment be disruptive to the policy, project, or program itself? [71]
	▪ Would an evaluation spark political reaction that could undermine the intervention by calling attention to it or by inadvertently feeding political forces in opposition to it? [72]

Developing TORs	
Implementers:	▪ What ethnic, religious, and linguistic groups do we need to include in the evaluation team to ensure a team that is reflective of those party to and affected by the conflict?
	▪ How substantive was the intervention's original and updated conflict analysis? Will the evaluator need to carry out extensive analysis or can they build off of existing materials?
	▪ Have we budgeted enough time for conflict analysis processes? What additional research resources might the evaluation team need if they are to carry out a sufficiently rigorous conflict analysis?
	▪ What areas of expertise (evaluation methods, in-depth experience with the context and conflict, experience with the funder/commissioner, etc.) are most crucial for this particular evaluative exercise?
Commissioners/ Implementers:	▪ Have we budgeted resources to adequately include the evaluation team in the planned processes for disseminating findings, to ensure that their findings are reported as they intended?

---

[71] OECD Evaluating Peacebuilding, 50
[72] OECD Evaluating Peacebuilding, 50

Inception	
Evaluator:	▪ What inherent values and biases could affect the way that we frame the evaluation and interpret findings?
	▪ Were the intervention's original and updated conflict analyses adequate?
	▪ Have we allotted significant time to enable training of TPM-affiliated enumerators and in-depth discussions on research ethics and the objectives of the research they will undertake?
	▪ Do the evaluand, evaluation commissioners and any other primary users agree on the evaluation's purpose and share a common understanding of the implications of potential negative findings? If not, how can we bring about a shared understanding?
	▪ What other steps can we take to limit the potential for interference during the reporting and dissemination stages?
Commissioner, Implementer and Evaluator:	▪ What is happening in the situation of conflict and fragility? At what stage is the conflict cycle? Watching carefully for potential conflict triggers (elections, controversial celebrations, etc.) [73]
	▪ Would an evaluation put stakeholders at personal or political risk? Will there be sufficient access to stakeholders, or sufficient safeguards, to avoid bias that might endanger the policy, program or project and the staff and stakeholders?[74]

Evaluation design	
Implementer and Evaluator:	▪ Will extensive consultations in the target communities raise expectations for aid and other interventions in those areas?
	▪ What effect might unrealistic expectations have on drivers of conflict?
	▪ What permissions are needed to access the geographical areas we seek to include?
	▪ How might obtaining those permissions affect perceptions of legitimacy of the civil or armed authorities in each area?
	▪ How might our inability to access certain areas or populations, evaluations in conflict zones cause our findings to be "biased, incomplete, and/or (voluntarily or involuntarily) censored"?
Evaluator:	▪ If seeking to involve vulnerable populations in research: Is the information that we are seeking available or accessible through some alternative means?
	▪ Does our intended use for the information justify the risk of re-traumatization of vulnerable populations?
	▪ What secondary information and local expertise can we draw on to ensure that our data collection processes do not cause additional harm?
	▪ Have we explicitly solicited the field team's input on the feasibility and safety of both accessing research sites and carrying out research protocols and on the appropriateness of each individual question?
	▪ Have we timed the data collection and findings dissemination processes in a manner that is sensitive to key dates/events?

[73] OECD Evaluating Peacebuilding, 50
[74] OECD Evaluating Peacebuilding, 50

	■ If submitting evaluation design for review by an IRB or ERB: Have you allowed enough time to adapt the design in response to feedback from the review board? ■ Who are we including as research participants? How might that choice affect conflict dynamics? Can we broaden the pool of participants to mitigate perceived favoring or marginalization of particular groups and sub-groups?
Developing data collection and analysis methods	
Evaluator:	■ How might the framing and wording of survey, interviews, and focus group questions exacerbate tensions?
	■ Are survey and interview questionnaires structured in a manner that protects respondents from having to associate themselves with information that could put them at risk (i.e. using list experiments, etc.)?
	■ Have we developed enough open-ended questions to access information about unintended consequences and prompt for questions outside the frame of the assumptions that underpin the program and evaluation designs? ■ How do local populations view the use of technology? Will providing data collectors with smart phones or tablets cause suspicion about their roles and intentions? ■ Who will analyze the data? What biases and assumptions do they bring to the process? ■ How can we both standardize data analysis processes to ensure adequate rigor and allow for in-depth sharing across perspectives that is necessary for conflict sensitivity? ■ What concerns were identified while piloting the data collection instruments? What subsequent modifications are necessary?
Data collection	
Evaluator:	■ Are the individuals who will interact with research participants fully trained on ethics? Have they practiced scenarios in which they need to modify planned protocols to protect participants from harm or re-traumatization? ■ *If working with TPMs:* Have we debriefed enumerators to gain insights into the tenor of their interactions with research participants and challenges or unexpected happenings during data collection? ■ *Examined continuously:* What changes have occurred that may affect the feasibility and safety of both accessing research sites and carrying out research protocols and on the appropriateness of each individual question?

Bibliography

Barnett, Chris and Jon Bennett. "Critical reflections on the South Sudan evaluation of conflict prevention and peacebuilding activities." In *Evaluation Methodologies for Aid in Conflict*, edited by Ole Winckler Andersen, Beate Bull, and Megan Kennedy-Chouane, 38–60. Oxford: Routledge, 2014.

Benelli, Prisca and Tamara Low. "The Ethics of Doing Primary Research with Vulnerable Populations: Facing the Challenges One Step at a Time." Paper submitted to Forced Migration Review, February 2019.

Broegaard, Eva, Beate Bull and Jens Kovsted. "Evaluation approaches in situations of conflict and fragility." In *Evaluation Methodologies for Aid in Conflict*, edited by Ole Winckler Andersen, Beate Bull, and Megan Kennedy-Chouane, 15-37. Oxford: Routledge, 2014.

Brusset, Emery, Marina Buch Kristensen, and Ndeye Bate Cissokho. "Evaluation of the USAID Peace-Building Program Casamance and Sub-region." *Channel Research* (2006): 5-61.

Brusset, Emery and Ivo Hooghe. "The case of Congo: an evaluation approach focusing on context." In *Evaluation Methodologies for Aid in Conflict*, edited by Ole Winckler Andersen, Beate Bull, and Megan Kennedy-Chouane, 85-104. Oxford: Routledge, 2014.

Bush, Kenneth, and Colleen Duggan. "Evaluation in Conflict Zones: Methodological and Ethical Challenges." *Journal of Peacebuilding & Development* 8 (2013): 5-25.

Bush, Kenneth and Colleen Duggan. "Research, Impact and Politics in Violently Divided Societies: An Evaluative Lens for Small Scale Maps." In *Evaluation in the Extreme: Research, Impact and Politics in Violently Divided Societies*, edited by Kenneth Bush and Colleen Duggan, 3-32. Thousand Oaks: SAGE Publications, 2015.

Campbell, Susanna P. "Literature Review: Ethics of Research in Conflict and Post-Conflict Environments." *Program on States and Security, The Graduate Center, City University of New York* (2010): 1-11.

Capable Partners Program (CAP). "Conflict-Sensitive Monitoring and Evaluation: Keys for Successful Programming." *Capable Partners Program* (Date unknown): 1-4.

Conflict Sensitivity Consortium. "How to Guide to Conflict Sensitivity." *Conflict Sensitivity Consortium* (2012): 4-46.

Corlazzoli, Vanessa, and Jonathan White. "Back to Basics: A Compilation of Best Practices in Design, Monitoring & Evaluation in Fragile and Conflict-affected Environments." *DFID Results Initiative in Conflict, Crime, and Violence Programming* (2013): 6-65.

Duggan, Colleen, and Kenneth Bush. "Lessons for Researchers and Evaluators Working in the Extreme." In *Evaluation in the Extreme: Research, Impact and Politics in Violently Divided Societies*, edited by Kenneth Bush and Colleen Duggan, 299-307. Thousand Oaks: SAGE Publications, 2015.

Ford, Nathan, Edward J Mills, Rony Zachariah, and Ross Upshur. "Ethics of conducting research in conflict settings." Conflict and Health 3:7 (2009): 1-9.

Gaarder, Marie, and Jennie Annan. "Impact evaluation for peacebuilding: challenging preconceptions." *In Evaluation methodologies for aid in conflict*, edited by Ole Winckler Andersen, Beate bull, Megan Kennedy-Chouane, Routledge, 2014. pp.130-154

Goodhand, Jonathan. "Research in conflict zones: ethics and accountability." *Forced Migration* (2000): 12-15.

Goodhand, Jonathan, Bart Klem, and Gunnar M. Sorbo. "Battlefields of method: evaluating Norwegian peace efforts in Sri Lanka." In *Evaluation Methodologies for Aid in Conflict*, edited by Ole Winckler Andersen, Beate Bull, and Megan Kennedy-Chouane, 61-84. Oxford: Routledge, 2014.

Gupta, Jhumka, Kathryn L Falb, Heidi Lehmann, Denise Kpebo, Ziming Xuan, Mazeda Hossain, Cathy Zimmerman, Charlotte Watts, and Jeannie Annan. "Gender norms and economic empowerment intervention to reduce intimate partner violence against women in rural Cote d'Ivoire: a randomized controlled pilot study." *BMC International Health and Human Rights* 13:46 (2013).

Hay, Katherine. "Building the Field of Evaluation in Violently Divided Societies." In *Evaluation in the Extreme: Research, Impact and Politics in Violently Divided Societies*, edited by Kenneth Bush and Colleen Duggan, 267-298. Thousand Oaks: SAGE Publications, 2015.

Herbert, Sian. "Remote management of projects in fragile states." *GSDRC, University of Birmingham* (2013): 1-9.

Humanitarian Outcomes Limited. "Summary Inception Report: Improving the Evidence Base on Delivering Aid in Highly Insecure Environments." Secure Access in Volatile Environments (SAVE) (2014): 4-64.

Humanitarian Outcomes Limited. "Interim Report: Component 3: Accountability and Learning." Secure Access in Volatile Environments (SAVE) (2015): 4-64.

Jayawickrama, Janaka, and Jacqueline Strecker. "The Ethics of Evaluating Research: Views from the Field." In *Evaluation in the Extreme: Research, Impact and Politics in Violently Divided Societies*, edited by Kenneth Bush and Colleen Duggan, 127-164. Thousand Oaks: SAGE Publications, 2015.

Mazurana, Dyan, and Lacey Andrews Gale. "Preparing for Research in Active Conflict Zones: Practical Considerations for Personal Safety." In *Research Methods in Conflict Settings*, edited by Dyan Mazurana, Karen Jacobsen, and Lacey Andrews Gale, 279-291. Cambridge: Cambridge University Press, 2014.

McDermott, Philip, Zahbia Yousuf, Jacqueline Strecker, and Ethel Mendez. "Fundamental Issues in Evaluation and Research in Violently Divided Societies: An Analysis of the Literature." In *Evaluation in the Extreme: Research, Impact and Politics in Violently Divided Societies*, edited by Kenneth Bush and Colleen Duggan, 33-92. Thousand Oaks: SAGE Publications, 2015.

Morris, Michael. *Evaluation Ethics for Best Practice: Cases and Commentaries*. New York: Guilford, 2007.

MSF. "Medecins Sans Frontieres: Research Ethics Framework: Guidance Document." *MSF ERB* (2013).

Ober, Heidi. "The What and Why of Third-Party Monitoring." *The Guardian*, May 6, 2015.

OECD. "Encouraging Effective Evaluation of Conflict Prevention and Peacebuilding Activities: Toward DAC Guidance." *OECD Journal of Development* 8:3 (2007): 7-97.

OECD. "Evaluating Peacebuilding Activities in Settings of Conflict and Fragility." OECD Publishing: *DAC Guidelines and Reference Series* (2012): 7-60.

OFDA. "Conflict Sensitive M&E Methods and Tools in Complex Humanitarian Emergences (CHEs): A Working Meeting for Practitioners, Program Managers and Experts." OFDA Training Facility: *Minutes from Working Meeting* (2006): 1-19.

Raftree, Linda, and Michael Bamberger. "Emerging Opportunities: Monitoring and Evaluation in a Tech-Enabled World." *The Rockefeller Foundation Evaluation Office* (2014): 1-46.

Richards, Simon, Sarah Bayne, Daniel Kiptugen, Terrence Lyons, Winnie Wambua, Manasseh Wepundi, and Management Systems International (MSI). "Performance Evaluation of USAID/Kenya's Conflict Mitigation and Civil Society Strengthening Activities." *Management Systems International* (2014): 1-45.

Rivas, Althea-Maria, David Guillemois, Kathryn Rzeszut, and Bradley Lineker. "Cross Cutting Evaluation of DFID's Approach to Remote Management in Somalia and North-East Kenya." *Integrity Research and Consultancy* (2015): 1-40.

Saferworld. "Chapter 3, Module 3: Conflict-sensitive monitoring and evaluation." In *Conflict-sensitive approaches to development, humanitarian assistance and peace building: tools for peace and conflict impact assessment* (2004).

Scharbatke-Church, Cheyanne. "Peacebuilding Evaluation: Not Yet All It Could Be." In *Berghof Handbook for Conflict Transformation*, edited by Beatrix Austin, Martina Fischer, and Hans J. Giessmann, 460-480. Berlin: Berghof Foundation, 2011.

Schultz, Christina. "Remote management is key to keeping aid workers safe in conflict areas." *Devex*, August 24, 2015.

Turner, David. "Evaluation Ethics and Quality: Results of a Survey of Australasian Evaluation Society Members." *AES Ethics Committee* (2003): 1-5.

USAID. "Conflict-Sensitive Programming Through Monitoring and Evaluation: A Report from USAID's Conflict-Sensitive Monitoring and Evaluation Workshop." *Capable Partners Program (CAP)* (2004): 1-23.

Villanueva, Starjoan D. "Managing Performance in Peacebuilding: Framework for Conflict-Sensitive Monitoring and Evaluation." *GoP-UN Act for Peace Programme* (2009): 7-33.

Wood, Elisabeth Jean. "The Ethical Challenges of Field Research in Conflict Zones." *Qualitative Sociology* 29 (2006): 373-386.