DAC Guidelines and Reference Series

Evaluating Peacebuilding Activities in Settings of Conflict and Fragility

IMPROVING LEARNING FOR RESULTS

OECD

DAC GUIDELINES AND REFERENCE SERIES

Evaluating Peacebuilding Activities in Settings of Conflict and Fragility

IMPROVING LEARNING FOR RESULTS

OECD
BETTER POLICIES FOR BETTER LIVES

This work is published on the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Please cite this publication as:
OECD (2012), *Evaluating Peacebuilding Activities in Settings of Conflict and Fragility: Improving Learning for Results*, DAC Guidelines and References Series, OECD Publishing.
http://dx.doi.org/10.1787/9789264106802-en

ISBN 978-92-64-10679-6 (print)
ISBN 978-92-64-10680-2 (PDF)

Series: DAC Guidelines and References Series
ISBN 1990-0996 (print)
ISBN 1990-0988 (PDF)

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Photo credits: Cover: © Holly Kuchera/Dreamstime.com.

Corrigenda to OECD publications may be found on line at: *www.oecd.org/publishing/corrigenda*.

© OECD 2012

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgement of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to *rights@oecd.org*. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at *info@copyright.com* or the Centre français d'exploitation du droit de copie (CFC) at *contact@cfcopies.com*.

Foreword

The international community, including members of the Development Assistance Committee (DAC), has paid increasing attention to situations of conflict and fragility, acknowledging that these settings represent some of the great development challenges of our time. Rising levels of resources go into these contexts, but the fact that no fragile state has yet to reach any of the Millennium Development Goals is a stark reminder to us all that results are difficult to achieve and sustain in these situations. Finding answers to improve delivery is urgent, not least for the populations suffering from conflict and poverty.

To deliver better results in situations of conflict and fragility we need to improve our understanding of the impacts and effectiveness both of programmes and projects aimed at supporting peace and of development and humanitarian activities operating in conflict settings. While the use of evaluation has become widespread in development and methods continue to evolve, it has grown clear that a special approach is needed to support learning and accountability in the context of conflict and fragility. How can evaluations provide strong evidence and lessons about what works and why in complex conflict settings – where change processes are non-linear and engagement politicised, and, where data are often missing or unreliable? How can the findings of these evaluations be used to inform policy making, programme design and implementation?

It is against this backdrop that I am pleased to present this OECD-DAC guidance on Evaluating Peacebuilding Activities in Settings of Conflict and Fragility and to highlight the timely and relevant contribution it makes to international policy debates. Whereas a few years ago it was thought that evaluation in situations of conflict and fragility was impossible – or even objectionable – the process of developing and testing this guidance has proven otherwise. Importantly, it has stimulated critical thinking and shown how an evaluative perspective can be useful to policy makers and practitioners – not just for commissioning evaluations, but throughout the programme cycle. It has also demonstrated that more rigorous assessment of the theories underlying donor action in settings of conflict and fragility can help debunk outdated myths about the role of aid in preventing violent conflict and supporting long-term development processes.

Evaluating Peacebuilding Activities in Settings of Conflict and Fragility was born out of a collaborative effort between evaluation and the conflict-and-fragile-state communities. By establishing a common understanding of key concepts and encouraging more and better evaluation, particularly of development co-operation activities in settings of conflict and fragility, it has created a community of practice bridging these diverse disciplines. I think it safe to say that the process itself has shown the value-added of cross-DAC work.

I encourage all those concerned with supporting positive change and sustainable development in today's fragile and conflict-affected regions to utilise this guidance – not just for commissioning evaluations, but as an input to learning and accountability throughout government.

J. Brian Atwood, Chair
OECD Development Assistance Committee

Acknowledgements

This guidance document has benefited from the valuable contributions of policy makers, practitioners, researchers, and conflict and evaluation experts, in development agencies and elsewhere, who want to see a strengthened evidence base and better results in settings of fragility and conflict. The OECD Development Assistance Committee (DAC) International Network on Conflict and Fragility (INCAF) and the DAC Network on Development Evaluation members and respective secretariats have led this collaborative process. The list of contributors over the years is long but a few should be specially mentioned for their role in completing the work.

Following the application phase, revision of the guidance was led by a joint task force made up of: Beate Bull of Norad's Evaluation Department; Ivo Hooghe from the Office of the Special Evaluator of International Co-operation in Belgium; Megan G. Kennedy-Chouane, of the OECD DAC Network for Development Evaluation Secretariat; and Asbjørn Wee of the OECD DAC INCAF Secretariat. Megan G. Kennedy-Chouane managed the process for the OECD. Special thanks are due to the Peer Review Panel for their review of early drafts: Ted Kliest from the Policy and Operations Evaluation Department (IOB) in the Netherlands Ministry of Foreign Affairs and Monika Egger Kissling from the Human Security Division, Directorate of Political Affairs, in Switzerland's Federal Department of Foreign Affairs. Special thanks also go to Norway and Norad's Evaluation Department for strong support and engagement throughout the process of developing the guidance, and for hosting the feedback workshop in February 2011. Mariska van Beijnum of the Clingendael Institute carried out useful revisions based on lessons from the application phase. Special appreciation goes to all those who guided the application phase and provided substantive input to this work over the years – especially Asbjørn Eidhammer, Cristina Hoyos, Dominique de Crombrugghe, Gørild Mathisen, Peter Woodrow, Diana Chigas, Anna Hellström, Hans Lundgren, Lisa Williams, Annette Brown, Marie Gaarder and Thania Paffenholz.

And finally, thanks to Ken Kincaid for editing, and to Jenny Gallelli, Alexandra Chevalier, Stephanie Coic and the OECD Public Affairs and Communication directorate for preparing this publication.

Table of contents

Executive summary 7

Glossary 11

Introduction 17
Why guidance on evaluating peacebuilding activities in settings of conflict and fragility? 17
Who will benefit from this guidance and how should it be used? 18
Scope and structure of the guidance 19

Chapter 1. **Conceptual background and the need for improved approaches in situations of conflict and fragility** 21
The need to better understand and adapt to conflict and fragility 22
Principles and objectives of peacebuilding and statebuilding support 23
Aid that does harm 26
Improving programme design and strategic planning 29

Chapter 2. **Addressing challenges of evaluation in situations of conflict and fragility** 31
Challenges to evaluations in situations of conflict and fragility 32
Overcoming challenges to evaluation 34

Chapter 3. **Preparing an evaluation in situations of conflict and fragility** 39
Defining the purpose 40
Analysing conflict 42
Deciding the scope of the evaluation 45
Selecting evaluation criteria 46
Outlining key evaluation questions 46
Incorporating gender equality and women's empowerment 48
Looking at the big picture 49
Selecting the best-fit evaluation methodology 49
Dealing with timing and logistics 50
Co-ordinating with other actors 51
Working with local and country stakeholders 52
Considering a joint evaluation 53
Writing terms of reference 53
Setting up evaluation management 54
Selecting the evaluation team 54
Controlling quality 54

Chapter 4. **Conducting an evaluation in situations of conflict and fragility** 57
Allow an inception phase 58
Identify and assess the theory of change and implementation logic 58
Gather data 61
Criteria for evaluating interventions 65
Draw conclusions and make recommendations 72
Reporting 73
Management response and follow-up action 74
Disseminate findings 75
Feed back into programming and engage in learning 75

Annex A. **Conflict analysis and its use in evaluation** 77
Annex B. **Understanding and evaluating theories of change** 80
Annex C. **Sample terms of reference for a conflict evaluation** 87

Bibliography 91

Tables

3.1. Some key questions for conflict analysis 45
3.2. Examples of evaluation scopes 45
A.1. Summary of selected conflict analysis tools 79
B.1. Common theories of change 85

Figure

4.1. Assumptions underlying the theory of change in a fictitious peace journalism programme 60

© OECD 2012

Executive summary

The high human, economic, political and social costs of violent conflict – coupled with a growing sense that such suffering and devastation could be avoided or at least mitigated – have led to increasing shares of development and humanitarian assistance being spent in settings of violent conflict and state fragility. In the decade to 2009 the share of overseas development assistance (ODA) to fragile, conflict-afflicted countries doubled to USD 46 billion and 37% of total available ODA. International actors now recognise the centrality of these challenges for global development.

Yet the scale of that effort is not reflected in its results. Findings from evaluations in these fields show that there are substantial weaknesses in programme design, effectiveness, and management. Peacebuilding and statebuilding support is often not based on a clear, strategic understanding of the conflict and (potential) role of international support in transforming key conflict drivers. Programmes lack basic conflict sensitivity and are not well adapted to the context in which they operate. The logic and assumptions underlying many activities in these fields are untested and objectives are unclear. Sketchy understanding of a conflict and unchecked assumptions can produce interventions that actually worsen tensions and fuel the conflicts they seek to mitigate.

The need for more and better evaluation in conflict settings

Furthermore, a persistent evaluation gap (few or weak evaluations of peacebuilding and conflict prevention activities), and little to no evaluation activity in settings of violent conflict, has meant that there is often very little credible information about the effectiveness and results of such endeavours. Learning and accountability have been weak. Research and experience, including the testing of the draft guidance, have shown that evaluations in these fields tend to be weak in terms of data, methods and validity of findings. Fewer rigorous methods are used and questions of causality are often inadequately addressed. Many evaluations in this field focus on process and mapping the context. Both internal and external validity tend to be quite low – meaning it is hard to draw broader lessons that can be applied to other contexts and it is difficult to draw credible conclusions about effectiveness and what works.

How and why guidance was developed

Evaluating Peacebuilding Activities in Settings of Conflict and Fragility (hereafter referred to as "the *Guidance*") was developed by the OECD's Development Assistance Committee (DAC) through the collaboration of its subsidiary bodies working on conflict and fragility and on evaluation. In 2008, the OECD produced a draft guidance, (*Guidance on Evaluating Conflict Prevention and Peacebuilding Activities – Working Draft for Application Period*) that was used to evaluate a range of activities. The findings from that application phase led to the 2008 draft

being revised. The result is the present guidance, *Evaluating Peacebuilding Activities in Settings of Conflict and Fragility*.

This guidance aims to help improve programme design and management and strengthen the use of evaluation in order to enhance the quality of conflict prevention and peacebuilding work. It seeks to guide policy makers and country partners, field and programme officers, evaluators and other stakeholders engaged in settings of conflict and fragility by supporting a better, shared understanding of the role and utility of evaluations, outlining key dimensions of planning for them, setting them up, and carrying them out.

This guidance is to be used for assessing activities (policies, programmes, strategies or projects) in settings of violent conflict or state fragility, such as peacebuilding and conflict prevention work and development and humanitarian activities that may or may not have specific peace-related objectives. This encompasses the work of local, national, regional and non-governmental actors, in addition to development co-operation activities. The central principles and concepts in this guidance, including conflict sensitivity and the importance of understanding and testing underlying theories about what is being done and why, are applicable to a range of actors.

Understanding key concepts

The document begins with a discussion of key concepts and provides an overview of current policy debates. It describes the convergence of the concepts of peacebuilding, statebuilding and conflict prevention and addresses the emerging international consensus that such contexts require specific, adapted approaches. It considers the principles for engagement in fragile states as the backdrop to evaluating such engagement and outlines the preconditions for evaluability, which should be handled by those designing and managing such programmes. Such conditions include setting clear, measurable objectives for peace-related activities, collecting baselines data and monitoring activities.

Challenges of evaluating in fragile, conflicted-affected settings

While no two situations of conflict and fragility are alike – all are specific to a place and time – they share some characteristics, many of which make evaluation particularly challenging. Evaluations in the field of conflict prevention and peacebuilding expose both evaluators and evaluated to violence. International engagement in settings of violent conflict is often highly politicised, complex and multifaceted, encompassing not only development activities, but ones that are humanitarian, diplomatic, and even military in nature. Evaluators may struggle to maintain safe "evaluation space" where they can produce credible, defensible findings. It can be particularly difficult to establish clear attribution and causality in settings that are complex and where changes for peace (or renewed violence) are often non-linear and unpredictable. Further complicating evaluation in these settings are the relatively weak programme designs and the lack of agreed upon, proven strategies for effectively working towards peace. Baseline and monitoring data, including information on implementation, is often lacking. These weaknesses in design and programme management can make peacebuilding and statebuilding activities less effective and particularly difficult to evaluate.

Overcoming challenges by understanding the conflict context

Moving from the discussion of challenges, the guidance outlines a number of basic principles to help overcome them, setting the stage for a description of key elements of planning for, preparing, implementing, and learning from an evaluation. The guidance argues that conflict analysis – which includes understanding conflict dynamics and actors as well as the economic and political context – is essential for designing and implementing strategies and programmes, as well as for evaluating such work. A clear analysis of the causes, drivers and dynamics of conflict and fragility sets the analytical framework for evaluation, and should also be used to ensure conflict sensitivity. Adjusting the evaluation to be sensitive to the conflict context may have implications for methodology, data collection and findings.

Conflict sensitivity and theories of change

Conflict sensitivity is an overarching principle highlighted in this guidance. All engagement (including evaluation) in such settings should be sensitive to conflict and avoid doing harm. However, conflict sensitivity does not of itself build peace and being sensitive to the conflict is not synonymous with being effective. Questions of conflict sensitivity will therefore be evaluated alongside an assessment of effectiveness and other criteria.

The concept of theory of change is presented as a way of encouraging critical thinking about the assumptions and strategies of peacebuilding and statebuilding. A theory of change is the understanding of how a specific activity will result in achieving desired changes in a particular context – it is the logic that underlies action. Developing better founded, more clearly stated theories about how peacebuilding and statebuilding can be achieved and supported is a key message from this guidance for decision makers, managers, and programme staff. Policies and programmes should use theories and assumptions that are tested and evidence based, which set out in clear cause-effect terms how they intend to produce outputs, outcomes and impacts. Doing so will not only help in the assessment of effectiveness and impact, but also contribute to knowledge about violence, peace and development. Untested or incorrect theories of change are often one reason why development assistance is failing to produce peacebuilding results. Evaluation contributes to testing theories of change and to building up the evidence base on peacebuilding and statebuilding.

Evaluations at work

Analysis against the DAC criteria constitutes the main substance of the evaluation study. Evaluators examine relevance, sustainability, effectiveness, efficiency and impact of activities in relation to the specific conflict context in order to answer the main evaluation questions posed in the terms of reference. The guidance describes how these criteria might be adjusted and offers examples of conflict-related lines of query. Data availability and other challenges may affect analysis, particularly when assessing impact. The last phase of an evaluation is to draw the conclusions and feed the findings into relevant planning, management, learning, research, or accountability processes. Dissemination strategies should be tailored to the target audiences, reaching them with timely, relevant information backed up by sound evidence. Actionable recommendations based on the conclusions should be presented as opportunities for learning and commissioning institutions should

ensure systematic response to the findings. Such an approach will increase receptivity and the chances that findings will be fed back into programme design and decision-making.

In these ways, more and better evaluation will contribute to identifying strategies and programmes that progress towards "peace writ large".

Glossary

Activity Actions taken or work performed through which inputs, such as funds, technical assistance and other types of resources, are mobilised to produce specific outputs (OECD, 2002). In this guidance, "activity" may include projects, programmes, policies and country assistance strategies. See also **intervention**.

Attribution The ascribing of a causal link to observed (or expected to be observed) changes and a specific intervention (OECD, 2002).

Baseline study An analysis describing the situation prior to a development intervention, against which progress can be assessed or comparisons made (OECD, 2002).

Conflict analysis A systematic study of the political, economic, social, historical and cultural factors that directly influence the shape, dynamics and direction of existing or potential conflicts. It includes an analysis of conflict causes and dynamics as well as assessments of the profiles, motivations, objectives and resources of conflict protagonists (CDA, 2007; Conflict Sensitivity Consortium, 2004).

Conflict mapping A representation of the main aspects of a conflict analysis, illustrating relationships between actors, causes, causal relationships, etc.

Conflict prevention Actions undertaken to reduce tensions and to prevent the outbreak or recurrence of violent conflict. Beyond short-term actions, it includes the notion of long-term engagement. It consists of **operational prevention**, (i.e. immediate measures applicable in the face of crisis), and **structural prevention**, i.e. measures to ensure that crises do not arise in the first place or, if they do, that they do not recur (OECD, 2001b; United Nations, 2001a).

Conflict sensitivity Systematically taking into account both the positive and negative impacts of interventions, in terms of conflict or peace dynamics, on the contexts in which they are undertaken, and, conversely, the implications of these contexts for the design and implementation of interventions (Conflict Sensitivity Consortium, 2004).

Counterfactual The situation or condition which hypothetically may prevail for individuals, organisations, or groups were there no intervention, e.g. the war that would have occurred had a peacebuilding intervention not taken place.

Country programme	One or more donor's or agency's portfolio of interventions and the assistance strategy behind them, in a partner (recipient) country.
Driving factors of conflict	The trends, currents, causes or fundamental influences that affect a conflict and help determine its characteristics, direction and ultimate outcome.
Do no harm	Ways in which international humanitarian and development assistance given in conflict settings may be provided so that, rather than exacerbating and worsening the conflict, it helps local people disengage from fighting and develop systems for settling the problems which prompt conflict within their societies (CDA, 2004; OECD 2010d).
Evaluability	Extent to which an activity or programme can be evaluated in a reliable, credible fashion. Evaluability assessments call for the early review of a proposed activity or programme in order to ascertain whether its objectives are adequately defined and its results verifiable (OECD, 2002).
Evaluation	Evaluation refers to the process of determining merit, worth or value of an activity, policy or programme. It consists of the systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors (Scriven, 1991; OECD, 2002).
***Ex ante* evaluation**	Evaluation performed before the implementation phase of an intervention (OECD, 2002).
***Ex post* evaluation**	Evaluation of an intervention after it has been completed (OECD, 2002).
Formative evaluation	Evaluation intended to improve performance, most often conducted during the implementation phase of projects or programmes.
Fragility, fragile state, fragile situation	National, regional and local territories where the state (including central and local authorities) has weak capacity to carry out basic governance functions, and lacks the ability to develop mutually constructive relations with society. Fragile states are also more vulnerable to internal or external shocks such as economic crises or natural disasters (OECD 2011b). These situations tend to be characterised by poor governance, to be prone to violent conflict, and to show limited progress towards development. An aggregate of governance and security criteria, or of capacity, accountability and legitimacy criteria are usually used as measures of fragility.

Goal	The higher-order objective to which a development intervention is intended to contribute (OECD, 2002).
Impacts	Positive or negative, primary and secondary effects produced by an intervention, directly or indirectly, intended or unintended (OECD, 2002). Results that lie beyond immediate outcomes or sphere of an intervention and influence the intensity, shape or likelihood of a conflict.
Indicator	Quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of a development actor (OECD, 2002).
Inputs	The financial, human, and material resources used for the development intervention (OECD, 2002).
Intervention	A general term that refers to the subject of the evaluation and may refer to an activity, project, programme, strategy, policy, topic, sector, operational area, country strategy, institutional performance, etc. Examples are policy advice, projects, programmes (OECD, 2010c).
Joint evaluation	An evaluation involving more than one donor agency and/or country partner (OECD, 2002).
Logical framework (log frame)	Management tool used to improve the design of interventions, most often at the project level. It involves identifying strategic elements (inputs, outputs, outcomes, impact) and their causal relationships, indicators, and the assumptions or risks that may influence success and failure. It thus facilitates planning, execution and evaluation of development interventions.
Monitoring	A continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an intervention with information regarding the use of allocated funds, the extent of progress, the likely achievement of objectives and the obstacles that stand in the way of improved performance (OECD, 2002).
Objective (project or programme objective)	The intended physical, financial, institutional, social, environmental, or other results to which a project or programme is expected to contribute (OECD, 2002).
Outcome	The likely or achieved short-term and medium-term effects of an intervention's outputs (OECD, 2002).
Outputs	The products, capital goods, and services which result from a conflict prevention and peacebuilding intervention (OECD, 2002).
Participatory evaluation	Evaluation method in which representatives of agencies and stakeholders (including beneficiaries) work together in designing, carrying out, and interpreting an evaluation (OECD, 2002).

Peace analysis An assessment of the peacebuilding environment, including existing peace efforts, actors, de-escalating factors (reduce armed conflict or tensions), and connectors (Paffenholz and Reychler, 2007).

Peacebuilding Actions and policies "aimed at preventing the outbreak, the recurrence or continuation of armed conflict", encompassing "a wide range of political, developmental, humanitarian and human rights programmes and mechanisms", including "short and long term actions tailored to address the particular needs of societies sliding into conflict or emerging from it" (UN, 2001b). Includes long-term support to, and establishment of, viable political and socio-economic and cultural institutions capable of addressing the proximate and root causes of conflicts, as well as other initiatives aimed at creating the necessary conditions for sustained peace and stability (OECD, 2001b).

Policy coherence The systematic promotion of mutually reinforcing policy actions across government departments and agencies, creating synergies towards achieving the agreed objectives (OECD, 2001c). There are four dimensions of coherence: a) consistency between ends and means of a policy; b) consistency of policies and activities across government departments; c) consistency of policies and activities pursued by different actors; and d) alignment of policies, activities and processes between external actors and conflict affected or conflict prone countries (Picciotto and Weaving, 2006).

Product An evaluation product is the output of an evaluation process and may take different forms, including written or oral reports, visual presentations, community meeting or videos.

Programme theory (see also logical framework and theory of change) A programme theory is a hypothesis or model of how a programme is intended to produce intended results and the factors affecting or determining its success. A programme theory often combines a theory of change and an implementation model (Bamberger *et al.*, 2006).

Project/programme cycle management (PCM) Management approach that systematically follows the cycle of planning, implementing, monitoring and evaluating the intervention.

Reliability Reliability refers to the consistency or dependability of data and evaluation judgements, with reference to the quality of the instruments, procedures and analyses used to collect and analyse data. Evaluation information is reliable when repeated observations under similar conditions produce similar results. Reliability contributes to credibility that can be additionally enhanced through a transparent evaluation process (OECD, 1991).

Results based management (RBM)	A management strategy focusing on performance and achievement of outputs, outcomes and impacts.
Result	The output, outcome or impact (intended or unintended, positive and/or negative) of a development intervention.
Risk assessment/risk analysis	An analysis or an assessment of factors (called assumptions in the log frame) that affect or are likely to affect the successful achievement of an intervention's objectives. A detailed examination of the potential unwanted and negative consequences to human life, health, property, or the environment posed by an intervention; a systematic process for providing information regarding such undesirable consequences; the process of quantification of the probabilities and expected impacts for (OECD, 2002).
Stakeholders	Agencies, organisations, groups or individuals who have a direct or indirect interest in the intervention or its evaluation (OECD, 2002).
Statebuilding	Statebuilding is a term used to describe the construction of a legitimate, functioning state. The OECD/DAC has defined statebuilding as an internal process to enhance capacity, institutions and legitimacy of the state, driven by state-society relations (OECD, 2008a).
Summative evaluation	A study conducted at the end of an intervention (or a phase of that intervention) to determine the extent to which anticipated outcomes were produced. Summative evaluation is intended to provide information about the worth of the programme (OECD, 2002).
Terms of reference (TOR)	A written document presenting the purpose and scope of the evaluation, the methods to be used, the standard against which performance is to be assessed or analyses are to be conducted, the resource and time allocated, and reporting requirements (OECD, 2002).
Theory of change	The assumptions that link a programme's inputs and activities to the attainment of desired ends. A set of beliefs about how and why an initiative will work to change the conflict. It includes both implementation theory and programme theory (Weiss, 1995; Church and Rogers, 2006).
Theory-based evaluation	An evaluation that tracks the anticipated sequence of linkages from inputs and activities to outcomes and impacts (Weiss, 1995).
Triangulation	The use of multiple theories, methods and/or data sources to verify and substantiate an assessment. It is used to overcome the biases that come from unitary disciplines, single observers, self-interested informants, and partial methods (OECD, 2002; Weiss, 1995).
Validity	The extent to which data collection strategies and instruments measure what they purport to measure (OECD, 2002). "External validity" refers to the extent to which findings or conclusions from one evaluation/context are applicable and valid in another.

© OECD 2012

Introduction

Why guidance on evaluating peacebuilding activities in settings of conflict and fragility?

In recent years, the international community has paid increasing attention to situations of conflict and fragility, acknowledging that they are one of the great development challenges of our time. As growing shares of resources, time and energy are devoted to projects, programmes, and policy strategies for countries affected by conflict and fragility, more evidence of the effectiveness of these endeavours is essential. Donors, practitioners and developing country governments show mounting interest in learning more about what does and does not work, and why, and in improving understanding of what contributes positively to sustainable peace and development.

The project of developing guidance to strengthen evaluation and learning in these contexts began with the identification of a persistent evaluation gap (too few or weak evaluations of peacebuilding and conflict prevention activities). Development actors undertake little to no evaluation activity in settings of violent conflict and the peacebuilding and conflict prevention fields have been under-evaluated (OECD, 2007a). Part of the explanation for the lack of evaluation activity is that evaluating in these contexts presents unique challenges. This guidance considers that the main challenge specific to evaluations in fragile and conflict-affected settings is the threat of violence. Other challenges covered in this guidance are: complexity, weak theoretical foundations, data collection, attribution, a highly political environment and multiple actors and multiple agendas. Challenges are further discussed in Chapter 2.

The lack of attention to evaluation and the challenges described above have meant that there is little credible evidence of the effectiveness and results of peacebuilding and conflict prevention endeavours. Research and experience, including the testing of the draft guidance, have shown that evaluations in these fields tend to be weak in terms of data, methods and validity of findings. Fewer rigorous methods are used and questions of causality are often inadequately addressed. Many evaluations in this field focus on process and mapping the context (FAFO 2006). Validity, both internal and external, tends to be low-meaning it is hard to draw broader lessons that can be applied to other contexts and difficult to draw credible conclusions about effectiveness and what approaches work.

The process of developing *Evaluating Peacebuilding Activities in Settings of Conflict and Fragility* (also referred to as "the *Guidance*") was spurred by a recognition in the peace and conflict prevention community of the lack of solid information about the actual results of peacebuilding efforts. Recognising the need for better, more tailored approaches to evaluation in conflict settings, the OECD's Development Assistance Committee (DAC) launched an initiative to develop guidance on evaluating conflict prevention and

peacebuilding activities. The initiative brought together practitioners and policy makers from the International Network on Conflict and Fragility (then the DAC Network on Conflict, Peace and Development Co-operation) with evaluation experts from the DAC Network on Development Evaluation. The OECD (2008a) produced draft guidance in 2008 which was used to evaluate various conflict prevention activities and external peacebuilding and statebuilding support in a number of major conflict settings including the Democratic Republic of Congo, Sri Lanka, Southern Sudan, and Afghanistan. The guidance has been revised on the basis of the substantive and methodological findings from this application phase.

The goal of this guidance is to promote critical reflection. It aims 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. At the same time, it aims to assist policy makers and practitioners working on peacebuilding and statebuilding to better understand the role and utility of evaluation and grasp how an evaluation lens can help strengthen programme design and management. With these objectives in mind, the *Guidance* offers advice on aspects of evaluating donor engagement in conflict-affected and fragile situations that differ from evaluation in more stable environments. To provide a complete picture it also covers some steps that apply to all development evaluations.

Who will benefit from this guidance and how should it be used?

Different target audiences will benefit in different ways from this guidance. The primary audience includes policy staff, donors, field and desk officers in foreign service offices and development agencies, partner country governments, non-governmental and international organisations (NGOs), and United Nations (UN) organisations involved in commissioning or supporting evaluations in situations of conflict and fragility. Secondly, it targets evaluators and evaluation managers, including the evaluation departments of developing countries and development agencies. Evaluators will benefit by gaining a clearer view of what commissioners expect from their work. Given the diversity of the intended audience, some sections may be more relevant than others to individual readers.

Evaluating Peacebuilding Activities in Settings of Conflict and Fragility provides an overview of key concepts relevant to evaluation in conflict situations and fragile states. It can be read while designing a programme or developing a strategic policy, while commissioning or programming an evaluation, and during the planning and carrying out of a specific evaluation.

This is not a prescriptive instruction manual. Rather, it seeks to contribute to fostering thoughtful, critical approaches by highlighting and clarifying specific challenges for evaluation. It should be viewed as a living guidance that will continue to evolve as evaluation methodologies and peacebuilding practices improve. It outlines key steps and main points to consider at each stage in the evaluation process and suggests tools that may support that process. The information and advice it volunteers should be applied carefully, based on an evaluation's context and intended purpose. To that end, this guidance is designed to be practical and to respond to the particular challenges that characterise fragile, conflict-affected situations and which evaluations must address.

Scope and structure of the guidance

Evaluating Peacebuilding Activities in Settings of Conflict and Fragility builds on existing literature and experience in development agencies and countries affected by conflict and fragility. This includes the lessons learned during the two-year application phase of a draft version of this guidance, when the suggested approach was tested in evaluations of external support in conflict settings. The draft guidance was employed for evaluations of multi-donor engagement in Southern Sudan (Bennett *et al.,* 2010), Sri Lanka (Chapman *et al.,* 2009), and the Democratic Republic of Congo (Brusset *et al*., 2011), as well as single-donor evaluations of the Norwegian contribution to peace in Haiti (Norad, 2009), the Swedish Provincial Reconstruction Teams in Afghanistan (unpublished), the German Civil Peace Service Programme (Paffenholz, 2011), and the European Commission's peacebuilding portfolio (EC, 2011).

Chapter 1 outlines the conceptual background of international engagement in settings of conflict and fragility, including main donor policy commitments, and examines why better understanding of conflict and fragility matters in today's development context. It is of particular relevance to those with limited experience in the conflict and peace domains and presents the overarching concepts that guide and inform decision making and evaluation.

If Chapter 1 is the theory, Chapters 2-4 are the practice. They form the "hands-on" core of the guidance and will be useful for all readers, particularly those with limited evaluation background. These chapters also provide seasoned evaluators with further ideas drawn from experience. They are guidance for planning, managing, implementing, and learning from evaluation. Underlying the chapters is the importance of understanding that each evaluation differs in its scope and purpose. Methodologies can and should be tailored accordingly.

Chapter 2 describes challenges to evaluation in settings of conflict and fragility. It then considers the principles that should guide evaluation and help it rise to the challenges of a fragile, conflict-affected setting. It emphasises the importance of a conflict analysis for assessing an intervention and for ensuring that the evaluation itself is conflict sensitive. Evaluations should also seek to be ethically responsible and transparent about strengths and weaknesses.

Chapter 3 considers the key steps in preparing an evaluation. It looks at the stages of defining an evaluation's purpose and scope and conducting a conflict analysis. It then examines timing and logistics, co-ordination with other actors, management methods, and hiring evaluation teams.

Chapter 4 deals with conducting an evaluation – from performing initial research to identifying the logic behind the development intervention, plugging gaps in data, and using OECD evaluation criteria of relevance, effectiveness, impact, sustainability and efficiency to assess the activity. Finally, it gives advice on follow-up, learning from evaluations, and feeding the lessons back into programming.

The annexes provide additional detail to complement Chapters 1-4. Annex A goes into further detail on conflict analysis, looking at different approaches and the use of the analysis in evaluation. Annex B provides further detail on the concept and use of theories of change. Annex C considers how to draw up a terms of reference document, using the example of an imaginary peace journalism training course. An extensive bibliography provides references and resources for further reading.

© OECD 2012

Chapter 1

Conceptual background and the need for improved approaches in situations of conflict and fragility

Chapter 1 outlines the conceptual background to working in settings of conflict and fragility. Arguing that such settings require a deep understanding of context and conflict, the chapter first seeks to characterise fragile and conflict-affected situations. It then looks at the purpose and goals of external engagement and describes, based on recent evaluations, how development assistance sometimes misses its targets and can even "do harm" when international partners have not sufficiently understood and adapted to the real context-specific drivers of peace and conflict. It is suggested that better conflict analysis and clearer targeting, together with more explicit and tested theories of change and results-based management can contribute to improving the knowledge base for development assistance programmes and facilitate evaluation.

The need to better understand and adapt to conflict and fragility

Armed conflict has devastating effects on human life. People in fragile and conflict-affected situations are more than twice as likely to be undernourished and lack clean water as those in other developing countries (World Bank, 2011). Children are affected particularly badly: a child in a fragile state is twice as likely to die before the age of five and also less likely to be able to attend school (*ibid.*) Violence and state fragility are often characterised by systematic violations of fundamental human rights. The impacts of conflict on political, social and economic development are also profound. When violent conflict breaks out, development is derailed.

Acknowledging the fact that countries affected by repeated cycles of political and criminal violence represent a central challenge for global development, donors provide them with substantial amounts of aid. Official development assistance (ODA) to fragile and conflict-affected states has doubled over the past decade, reaching USD 46 billion in 2009 and accounting for 37% of the total available ODA (OECD, 2011d). There is, however, an increasing body of evidence to suggest that aid aimed at achieving sustainable peace and development is not making a lasting contribution to peace and development. In 2005, a review of more than 75 evaluations in the conflict fragility field pointed to substantial weaknesses in programme effectiveness, design, and management (Fafo Institute, 2006). These findings were confirmed during the application of the earlier draft of this guidance (Kennedy-Chouane, 2011).

In Southern Sudan for example, it was found the support provided by multiple donors in 2005-2010 was often mistargeted. Because donors did not fully take into account key drivers of violence, there was an overemphasis on basic services and a relative neglect of security, policing, and the rule of law, which were found to be essential in the process of state formation for the future South Sudan and therefore, critical to preventing future conflict (Bennett *et al.*, 2010). Similarly, in the Democratic Republic of Congo, it emerged that one of the principal conflict drivers is that the Congolese justice system lacks credibility, political commitment, and competence and maintains a delicate relationship with customary law. The justice system is particularly inept in dealing with complex land ownership conflicts, which fuels violence and human rights violations, particularly where populations have been displaced and in the context of the unregulated exploitation of natural resources. Several large-scale multi-donor projects have targeted the restoration of justice and the rule of law. However, issues relating specifically to property titles, rent, and land rights have not been treated adequately within these programmes, according to local government and community groups surveyed (Brusset *et al.*, 2011).

There is no universal definition of fragile and conflict-affected situations – analysts and donors still have different notions of exactly what is meant. The OECD *Principles for Good Engagement in Fragile States* (2007a) – known as the *Fragile States Principles* – outline that in fragile situations, governments lack the political will and/or capacity to fulfil the basic conditions for poverty reduction, development, security, and human rights. In other words, vicious cycles of conflict commence when political and economic stresses and pressures

on justice and security meet weak institutions (World Bank, 2011). The OECD (2011b) also states that fragile and conflict-affected states are those that have weak capacity for carrying out the basic functions of governing their populations and territory and lack the ability to develop mutually constructive and reinforcing relations with society.

As pointed out in the *Fragile States Principles*, fragile and conflict-affected situations cover a broad spectrum. However, they do share some common features:

- They are inherently high-risk environments – for the people who live there, for their governments, and for those who provide and implement humanitarian and development assistance. The risks are not only related to the security situation (e.g. threats to staff, difficulty of movement, lack of access to information), but to the achievement of development objectives: risk of programme and strategic failure, fiduciary risks (corruption), and risks to the reputations of donors and implementing agencies.
- They are characterised by complex political economies and state-society relations, in which development partners can be parties to on-going conflicts and contested peace processes complicate efforts to prioritise needs and identify a strategic vision for sustainable peace.
- They are most often characterised by weak or non-existent national and local capacities and institutions. They are thus incapable of identifying or building consensus on peacebuilding priorities, developing strategies, implementing programmes, or monitoring progress.
- They are exposed to a combination of internal and external stresses that heighten the risk of violent conflict. Internal causes of conflict arise from political, economic, social and security-related dynamics (e.g. political exclusion, legacies of violence, crime, low GDP per capita, unemployment, identity-based conflict, and inequality). External stresses and regional conflicts can further exacerbate internal stresses (price shocks, for example, impact on inequality and unemployment) and some, like drug trafficking, can even cause them (World Bank, 2011).
- External humanitarian and development action is often part of donor's broader geopolitical and economic agendas – such as combating international terrorism, stabilising access to scarce resources like oil, fighting transnational organised crime, opening markets for domestic firms and curbing immigration flows. As such, aid is at a higher risk of being politicised in fragile, conflicted situations than in more stable ones, and development actors may not be in the lead in setting the agendas for engagement.

Principles and objectives of peacebuilding and statebuilding support

The past few years have seen an increase not only in international engagement in situations of conflict and fragility, but also in the convergence of development, security, human rights, humanitarian assistance, peacebuilding, statebuilding, and related agendas. It is internationally acknowledged that sustainable peace and development are critically linked to the capacity and legitimacy of the state. Donors need to base their interventions not only on the need to support short-term stability (or the cessation of hostilities) or on the provision of humanitarian aid, but on a broader understanding of how their interventions affect state-society relations and longer term prospects for the development of a functioning, legitimate state. Donor engagement in fragile and conflict-affected states is largely guided by the overarching aims of preventing conflict, peacebuilding, and statebuilding (see Box 1.1).

Box 1.1. **Key donor agendas for situations of conflict and fragility**

Conflict Prevention

Conflict prevention refers not only to actions undertaken in the short term to reduce manifest tensions and to prevent the outbreak or recurrence of violent conflict (OECD; 1997, 2001a). It also includes long-term engagement that addresses the built-in capacities of societies to deal with conflicting interests without resorting to violence (Menkhaus, 2006), and extends to the management of disputes with destabilising potential. Such work helps de-legitimise the belief that violence is an inevitable or acceptable way of resolving disputes, making nonviolent alternatives known and more attractive, addressing structural and immediate causes, and reducing vulnerability to triggers. The goal is not to prevent **all** conflict. Some conflict is natural, inevitable, and a positive part of development and other change processes. Instead, the emphasis is on preventing harmful **violent** responses to the inevitably diverging interests and conflicting objectives that exist in all societies.

Peacebuilding

Although most peacebuilding focuses on the transition from war to peace, the concept and practices of peacebuilding are, in principle, about supporting sustainable peace, regardless of whether or not political conflicts have recently produced violence. Indeed, the mere threat of violence occurring is sometimes enough to kick-start a peacebuilding process. Peacebuilding, in other words, is undertaken because violent conflict is looming, is going on, or has recently ceased (OECD, 2011b). The emerging UN consensus (2007) is that:

> "Peacebuilding involves a range of measures aimed at reducing the risk of lapsing or relapsing into conflict, by strengthening national capacities for conflict management and laying the foundations for sustainable peace. It is a complex, long-term process aimed at creating the necessary conditions for positive and sustainable peace by addressing the deep-rooted structural causes of violent conflict in a comprehensive manner. Peacebuilding measures address core issues that affect the functioning of society and the state."

Such wording points to a preventive, as well as a post-conflict, role for the concept and practice of peacebuilding.

Statebuilding

Statebuilding has been defined by the OECD DAC as "an endogenous process to enhance capacity, institutions and legitimacy of the state driven by state-society relations" (OECD, 2008a). The process must be understood against a background of long-term historical and structural factors that contribute to shaping the contours of state formation and the nature of state-society relations. And it must be understood within the exigencies of current circumstances in the country concerned. These may include the risk of violent conflict or effects of previous conflict – either internally or in the region – or the impact of economic pressures generated by global recession, debt, limited trade opportunities, financial imbalances and commodity prices (OECD, 2011b).

The objectives of conflict prevention, peacebuilding, and statebuilding are inextricably linked. Efforts to support and achieve them essentially address the same underlying problems. Their aims, too, are consistent: to help societies move in directions conducive to nonviolent resolution of conflict, address grievances and injustice, and move towards

sustained peace and development (OECD, 2011b). The wished-for end result of donor engagement in situations of conflict and fragility is not simply the absence of open conflict but a deeper peace, often referred to as "peace writ large", i.e. societal-level peace or the bigger peace beyond the micro level of a single project (CDA Collaborative Learning Projects, 2004). The complicating factor is that peace and conflict are context-specific: there is no one blueprint either for the end state or the means of achieving it that can be applied to all situations.

Donors recognise that much remains to be done to improve their engagement. Working with partner countries, they have committed themselves to a number of principles and guidance documents that underline what differentiates engagement in fragile and conflict-affected situations (Box 1.2) from development co-operation in other settings.

Box 1.2. **Principles for donor engagement in situations of conflict and fragility**

- Principles for Good International Engagement in Fragile States and Situations (OECD, 2007a).
- Guidelines for actors involved in development co-operation, peacebuilding, statebuilding and security in fragile and conflict-affected states.
- Supporting Statebuilding in Situations of Conflict and Fragility: Policy Guidance (OECD, 2011b).
- Actionable guidance on the way development actors provide support to statebuilding in fragile and conflict-affected situations with a focus on strengthening state-society engagement.
- The New Deal for International Engagement in Fragile States (OECD, 2011c).
- Agreed at the 4th High Level Forum on Aid Effectiveness in Busan, South Korea, the New Deal aims to improve the effectiveness of aid in contexts of conflict and fragility. It sets out five peacebuilding and statebuilding goals and outlines how partners will work towards achieving them.
- International Support to Post-Conflict Transition: Rethinking Policy, Changing Practice (OECD, 2012).
- Presents recommendations for better practice in order to improve the speed, flexibility, predictability and risk management of aid during transition.
- Paris Declaration on Aid Effectiveness (2005), Accra Agenda for Action (2008) and the Busan Declaration on Effective Development Co-operation (2011).
- International commitments to improve effectiveness of development co-operation including by increasing co-ordination and country ownership, adapting to differing country environments and giving increased attention to fragile and conflict-affected countries.
- Managing Risk in Fragile and Transitional Contexts (OECD, 2011e). Provides information to help donors understand how to balance risks and opportunities in order to protect the integrity of their institutions while delivering better results to those who need it most.
- Principles and Good Practice of Humanitarian Donorship (GHD, 2003).
- United Nations Security Council Resolution 1325 (United Nations Security Council, 2000). Establishes that equality between men and women is essential to achieving and sustaining peace. Calls for the protection of women and girls and for equal participation in peace processes and post-conflict reconstruction efforts.

One particularly important international commitment came in 2007, when OECD country ministers approved the *Fragile States Principles* as a guide to donor engagement in fragile states. The *Principles* highlight the importance of viewing countries in their particular contexts and thinking carefully about the objectives and likely impact of specific activities. They also underline the peculiarities of fragile states, which call for well sequenced and prioritised action across political, economic, administrative, and security-related domains. Such an effort entails shared analysis, objectives, strategies and resources. The ten principles are:

1. Take context as the starting point.
2. Ensure all activities do no harm.
3. Focus on statebuilding as the central objective.
4. Prioritise prevention.
5. Recognise the links between political, security and development objectives.
6. Promote non-discrimination as a basis for inclusive, stable societies.
7. Align with local priorities in different ways and in different contexts.
8. Agree on practical co-ordination mechanisms between international actors.
9. Act fast… but stay engaged long enough to give success a chance.
10. Avoid pockets of exclusion ("aid orphans").

Recent work by the OECD (2011e) has also helped to build consensus among development partners on the need to better manage and mitigate risks. Experience shows that donors tend to focus on fiduciary risks and ones that jeopardise the reputations of development agencies, using them as a reason not to engage in high-risk conflict or post-conflict countries. Development partners should, however, think more about contextual risks – the re-emergence of violent conflict, humanitarian disasters, etc. – and accept that the risks of non-engagement are often higher than those of engagement (OECD, 2011c). Risk mitigation does not mean eliminating risk, but, rather, finding appropriate ways of dealing with it. Shared risk assessments can be one such way.

The international community has made much progress towards understanding – and improving – the role of external partners in settings of conflict and fragility. These *Fragile States Principles* have contributed to changing donor policy and, to some extent, donor behaviour (OECD, 2011d). However, the effects and results of applying them have not been rigorously evaluated.

Aid that does harm

There is an emerging understanding that ill-designed, poorly implemented, or badly co-ordinated interventions in fragile and conflict-affected situations can increase tensions and undermine capacities for peace. They can, in other words, "do harm" (Anderson, 1999a). Conflict sensitivity is needed to mitigate such harm by systematically taking into account both the positive and negative impacts of interventions (International Alert, 2007a). The implication for approaches used to deliver aid is that they need to be tailored to the demands of high-risk environments.

Donors need to be realistic about what they can achieve as external partners in limited timeframes with limited capacities. Too often they underestimate the challenges of engaging in fragile and conflict-affected situations and draw up plans and schedules that

have little grounding in reality. This tendency was confirmed by findings of the pilot evaluations carried out during the development of this guidance. Evaluations have also found that donors fail to prioritise their engagement and lack clear strategies to address core peace drivers and conflict-mitigating factors. A focus on providing humanitarian aid or basic services – and a neglect of key priorities aimed at building and sustaining peace – tends to be driven by the untested assumption that all development activities will somehow contribute to peace (Kennedy-Chouane, 2011). The illuminating example of a well-meaning intervention in Tajikistan (Box 1.3) demonstrates the critical importance of understanding the real post-war peace drivers of a particular context.

Box 1.3. **Do no harm – an example from Tajikistan**

At the end of the civil war in Tajikistan, one international NGO undertook massive housing reconstruction in a southern province. The intent of the effort was to i) to encourage people displaced during the fighting to return to the region and ii) to support reconciliation between the former foes by getting them to work together in rebuilding the destroyed villages. Priority for reconstruction went to the villages that had suffered the most damage. In these, the NGO worked with local people to decide which houses would be rebuilt and organise work crews to do the construction. They agreed that "anyone from the village who wanted a job" would be hired in these crews.

A few months later, they had successfully sponsored the reconstruction of almost 60% of the damaged housing in the region. However, one day a local man came into the NGO's compound with a Kalashnikov and threatened the staff, saying, "Why are you favouring that group that we defeated in the war? If you don't start building some houses for my clan, I will kill you."

The NGO staff members were astounded. They had meant to be completely inclusive and to ensure that everyone who suffered in the conflict received equal attention. What they had not known until this moment was that during the conflict, the greatest damage had been done in villages occupied by only one (rather than both) of the local warring groups. By focusing their assistance on the areas of greatest damage, and by hiring people from the villages in those areas to work on the rebuilding, they had inadvertently provided almost all their assistance to one side of the conflict – and the "losing side" at that. Their project design had unintentionally reinforced existing inter-group divisions by focusing on mono-ethnic villages and channelling all their support to one group.

With a project redesign the NGO was able to supply building materials and support to multi-ethnic villages, to the damaged homes of the other ethnicity, and to community buildings that both groups shared such as schools, clinics and mosques.

Source: CDA Collaborative for Development Projects (2000).

Another widespread assumption is that being "conflict sensitive" is, *ipso facto*, doing peacebuilding work (OECD and CDA, 2007). As a result, much of the ODA aimed at fragility and conflict does not affect their driving factors. And, even when it is better targeted, it often remains ineffective. A case in point described by Bennett *et al*. (2010) is Southern Sudan. Between 65% and 85% of the total aid, including humanitarian assistance, from multiple donors in 2005-2010 targeted traditional socio-economic aid sectors. While this aid was provided in a way that many partners described as "conflict sensitive" (avoiding exacerbating ethnic tensions and trying to right historic inequalities), the conflict analysis

Box 1.4. **Weaknesses around conflict analysis**

Experience shows a number of recurrent challenges to the production and use of conflict analysis.

- **Partial analysis**. Due to time or resource constraints, it is often tempting to limit the focus of a conflict analysis to a donor's particular programme or strategy and how it might fit the context. Such an approach can lead practitioners to miss important aspects of the context or to develop misguided or irrelevant programmes.
- **Many people carry out context analysis, believing it to be conflict analysis**. A context analysis seeks a broad understanding of the entire political, economic and social scene. A conflict analysis is more narrowly focused on the specific elements of that broader picture that may cause, trigger, or propel incompatible interests or violence. Conflict analysis focuses on those political, economic, social, historical and other factors that directly influence the shape and dynamics of the situation of conflict and fragility.
- **Analysis is not updated**. Analyses are performed only at the front end of a programme. There are seldom efforts at ongoing in-depth analysis or monitoring and adjusting over time.
- **Programming is not linked to analysis**. Even when practitioners do perform an analysis, they often fail to link their programme strategy to it or adjust activities and strategies to changing dynamics over time.

Many implementing agencies and field staff work on the basis of an implicit analysis, often based on their own experience. Some programmes – frequently effective ones – are grounded in an informal analysis that draws on the long experience of local people or long-time observers of a conflict. These analyses can be quite sophisticated and may be constantly updated as individuals move about and talk with many different people. However, when analysis is done in this way, different members of the same project team or organisation sometimes operate on the basis of quite different understandings of the situation and their programme's role in it. This undermines the development of coherent strategies, weakens sustainability (when staff leave, so does their analysis) and significant assumptions often remain undiscussed and untested. Therefore, efforts to make the implicit analysis more explicit and to share observations are usually valuable.

Source: CDA Collaborative Learning Projects (2009).

conducted by evaluators showed that lack of social services could be cited as neither the sole nor a significant cause of conflict. The aid, while avoiding doing harm, was clearly not addressing the sectors most likely to be factors in sustainable peace.

To contribute to sustainable peace, donors should work on different priorities across humanitarian, development, conflict prevention, stabilisation, and peacebuilding activities. Priorities may need to be adjusted over time as a conflict evolves or political context shifts. Such an approach often involves a combination of working "in" and "on" conflicts. There is a widely held belief that traditional development activities (in areas such as health and education) can have a positive impact on conflict dynamics. However, this assumption needs to be critically examined and there is a growing consensus that development work should be complemented by activities that focus specifically on removing the causes and drivers of conflicts and strengthening the capacities, institutions and norms necessary for conflict management. Evaluation experience shows that the main issue in determining the effectiveness of donor engagement in situations of conflict and

fragility is not the effect that activities labelled as "peacebuilding" have on peace. It is much more closely related to the influence that all forms of aid combined have on peace (Chapman *et al.*, 2009; Bennett *et al.*, 2010; Brusset *et al.*, 2011). Development assistance should deliberately work in and on conflicts rather than simply endeavouring to get round them (OECD; 1998, 2001b).

There is a real and growing need for thoughtful examination of development actors' policy and practice in countries affected by violent conflict and state fragility. Consensus does exist that distinctive approaches are required to deliver effective support. Yet more work is needed to operationalise conflict-sensitivity concepts and achieve actual change at donor headquarters and country level, especially in terms of knowledge of conflict dynamics and how interventions relate to conflict, increased emphasis on outcomes and impact, and better understanding of the means of achieving these. Further work is also required to understand how individual donors should engage with national governments and align with country priorities and systems in situations where state legitimacy is weak or states are actors in violent conflict. More and better evaluation will contribute over time to helping practitioners better understand how to make their interventions more conflict sensitive, as well as more effective.

Improving programme design and strategic planning

Because good programme design is key to not only working effectively in support of peace and development, but also a prerequisite for good evaluation, it is important to consider the basics of planning, monitoring, and management. Evaluation findings in recent years have shed light on core dimensions of quality programming. When evaluation and its requirements are an integral part of programming activities from the outset, it contributes to more effective programming and facilitates better evaluation. As the OECD (1991) states: "clear identification of the objectives which an aid activity is to achieve is an essential prerequisite for objective evaluation". Some basic components of good programming are baseline information on key indicators, conflict analysis, clear and measurable objectives, a testable programme logic or theory, and monitoring data. These elements of programme design and management may, for a variety of reasons, be lacking or missing from assistance activities in fragile and conflict-affected situations, especially ones where fragility is prolonged and there is open, armed conflict.

- **Planning and conflict analysis** entails identifying the most relevant contribution(s) that donors, governments and other actors can make to support peacebuilding and reduce fragility in a specific country or conflict setting. Concrete understanding of the political, economic and social dimensions of conflict and fragility is a necessity. It is good practice to conduct a context and conflict analysis as one of the first steps in planning. The understanding that is gained should influence strategy, policy, and programme design and implementation. (Box 1.4 describes some of the weaknesses of development strategies and activities related to conflict analysis.)
- **Setting clear, realistic objectives** with clear target outcomes related to the context deserve specific attention. Peacebuilding and statebuilding programmes and policy goals tend to be general, vague, and consequently difficult to manage and evaluate.
- **Theory of change** is a set of beliefs about how change happens and, as such, it explains why and how certain actions will produce the desired changes in a given context, at a given time (Weiss, 1995; Church and Rogers, 2006). Developing a sound, clear, evidence-based theory of

change is one potentially useful way to improve design. Theory of change thinking is an approach that encourages critical thinking throughout the programme cycle.

- **Results-based management** is a management strategy that focuses on performance and the achievement of outputs, outcomes, and impacts. A key component is monitoring, which tracks how a programme is progressing and enables the adjustment of activities and strategies as the conflict setting shifts. Effective monitoring plays a crucial role in making programmes flexible and adaptable to changing contexts, particularly in complex situations of fragility and violence.

In summary, reliable, comprehensive programme design, management, and monitoring can contribute to better policies and programmes and improve the effectiveness of interventions. They also set the stage for successfully evaluating conflict prevention and peacebuilding programmes, notably by creating a theoretical framework and setting up necessary data management systems. Programme planners, policy makers, implementing staff, and evaluators should work together to strategise about how best to develop a well designed, effectively monitored, evaluable intervention.

© OECD 2012

Chapter 2

Addressing challenges of evaluation in situations of conflict and fragility

This chapter is first of the three that form the main evaluation guidance. Building on the conceptual basis of Chapter 1, it outlines key challenges to evaluation in these settings and then describes core principles for addressing these challenges, including the OECD evaluation principles. The chapter considers the role of conflict analysis and the need to understand the particular context of the intervention. These principles should guide an evaluation in fragile, conflicted settings throughout the process described in Chapters 3 and 4.

This chapter describes some of the key challenges to evaluating in settings of conflict and fragility and then sets out the core principles for meeting these challenges.

Challenges to evaluations in situations of conflict and fragility

This guidance considers that the main challenge specific to evaluations in fragile and conflict-affected settings is understanding and adapting to violent conflict, while mitigating the risk that evaluations themselves become part of the conflict or cause harm to those involved. Other challenges addressed are: complexity, weak theoretical foundations, challenges to data collection, attribution, a highly political environment, multiple actors and multiple agendas.

The high risk of violence

Evaluations of interventions in the field of conflict prevention and peacebuilding expose – in contrast to almost all forms of evaluation – both evaluators and evaluated to real risk. Potential implications are profound. First, the threat of violence may constrain the evaluators' ability to raise issues, collect material and data, recruit and retain local staff, meet interlocutors, publish findings, and disclose sources. Defending the integrity of evaluation findings in highly politicised and even dangerous settings can pose problems for evaluation teams, particularly where evaluation findings may potentially be misused by different parties to a conflict or harm those involved. Second, the risk of harm may mean that the information obtained is biased, incomplete and/or (voluntarily or involuntarily) censored. Consequently, evaluations must address the operational and methodological consequences of the risk of violence. More specifically, in order to deal with this challenge, it is advisable that the evaluation itself include a conflict analysis in order to assess the intervention and to ensure that the evaluation process and product is conflict sensitive.

Complex and unpredictable contexts and interventions

Few would dispute that settings of conflict and fragility are complex, combining multifaceted, multi-directional change processes with high levels of unpredictability, a general lack of information, and potential strategic misinformation. The way programmes are implemented on the ground may differ widely from original plans, as practitioners change what they are doing to adapt to an evolving conflict. As a result, it may be difficult to identify what exactly should be evaluated. Although unpredictability and complexity may be inevitable, their frequently negative ramifications for evaluations need not be. Evaluators must prepare for risks, develop robust designs, and ensure sufficient flexibility to counter the challenges of unpredictability and complexity. They should select methods that help to capture complex social change processes and illuminate interactions between interventions and the context.

Multiple actors

Many players work in fragile and conflict-affected settings, seeking to effect change and influence the situation, which adds additional dimensions of complexity and uncertainty. Actors may be members of the diplomatic corps or the military; development and humanitarian agencies or government bodies; informal power structures or various local groups. These many actors have different cultures, loyalties, institutional features and interests, and do not always pull in the same direction. There are also differences in terminologies, planning cultures, and approaches between the different agencies on the donor side.

Weak theoretical foundations and evidence base

The theories underpinning international support to peacebuilding, conflict prevention and statebuilding are weak. There is a lack of agreed upon, proven strategies for effectively working towards peace. The logic underpinning donor activities is often unclear. Numerous strategies and programmes are poorly designed with ill-defined objectives and a lack of clearly stated, tested (or testable) theories of change (i.e. the implicit or explicit understandings of how it is hoped that what is being done will contribute to peace). Programme approaches are often contested and evolve rapidly to adapt to the changing context, meaning it may be difficult to establish what activities and strategies are actually being implemented. All of which makes programmes less easily "evaluable".

Challenges to data collection

Challenges encompass scarcity of data, lack of monitoring, high personnel turnover, and erratic access to field data in certain regions at certain points in time. While the lack of timely, relevant, comparable data of high quality is not unique to situations of conflict and fragility, data problems tend to be compounded in these settings due, for example, to weak state statistical capacities and a multiplicity of international actors with incoherent data systems. However, this guidance suggests that more for data collection sources are available than currently used and that resources and institutions with special competence in this area exist and should be taken advantage of.

Attribution

Attribution is the ascribing of a causal link from a specific intervention to observed (or expected) changes. While attribution poses a problem in all areas of development work, attributing results to any particular policy or single intervention in conflict contexts is even more difficult. The difficulty arises principally from the fluidity and complexity of conflicts settings themselves and from frequently non-linear nature of change processes. For example, other activities (beyond the scope of the evaluation), such as military interventions, may actually be responsible for changes that are attributed to conflict prevention or peacebuilding activities. It can be very difficult for evaluators to control for these outside variables. Related challenges include the difficulty of creating a counter-factual or control group, especially when looking at country or regional conflicts, which is necessary to describe with reasonable certainty what would have happened had the activity in question not taken place

Politicisation

Fragile and conflict-affected settings are highly political environments. Due to the politicisation of international involvement and political sensitivities in national contexts, evaluators may find it difficult to maintain a safe, credible "evaluation space".

Box 2.1. **Political constraints in conflict settings: lessons from Sri Lanka**

The evaluation took place in 2008-2009 in the complex political context of strained relations between the government and donors and significant security restrictions on travel outside Colombo. This limited the range of parties that could be interviewed and made donors hesitant to release sensitive strategy and programming material to the evaluation team. In response to the worsening security situation, the evaluation team decided to exclude from the sample the activities associated with the peace negotiations, as well as other aspects of diplomatic engagement, the security sector, and some donors' internal analyses. This was a significant decision that led to agreement by most donors to support the evaluation, even though it meant important areas of donor engagement and the history of conflict prevention and peacebuilding work were not assessed.

Source: Chapman *et al.* (2009).

Overcoming challenges to evaluation

This section outlines core principles for evaluation in settings of violent conflict and fragility. These principles should be carried throughout the evaluation process, informing each of the steps outlined in Chapters 3 and 4. The specific issues outlined here complement the general evaluation principles and standards which the OECD has set out. The application phase of this guidance showed that these general evaluation principles, as outlined below are also relevant and valid in conflict settings. There is no excuse not to apply them. When applied carefully, they enhance the credibility, use, and rigour of the evaluation process and its end results.

The OECD DAC's *Quality Standards for Development Evaluation* (2006a) provides a guide to good practice. This short document, built through international consensus, is a staple reference for all evaluations, including those in settings of conflict and fragility. The standards draw on the core principles that evaluation processes should be impartial, credible, transparent, and independent. They should also be useful and relevant, informing decision makers and contributing to learning. An evaluation report should describe transparently the data sources, data collection instruments and analytical methods used and identify their strengths and weaknesses. Evaluation teams should deal with attribution and causality in a credible way. Commissioning agencies should make the results of evaluation widely available and ensure that they are used systematically by decision makers and others to support learning and accountability. The standards also state that the collaboration of development partners is essential. This and other key references for evaluation are presented in OECD (2011d), *Evaluating Development Co-operation: Summary of Key Norms and Standards*.

Context as the starting point: conflict analysis

What is known about a situation of conflict and fragility, its causes, components, and dynamics? Conflict analysis – which includes analysis of the political economy, stakeholders,

and conflict drivers and causes – is central to any evaluation of donor engagement in situations of conflict and fragility. Conflict analysis provides an analytical framework for assessing the relevance, effectiveness, and impact of peacebuilding activities, as explored in Chapter 4.

Conflict analysis may be used as the basis for assessing whether activities have been sufficiently sensitive to the conflict setting, determining the scope of the evaluation (what will be evaluated), and identifying pertinent evaluation questions. Another of its functions is to ensure that the evaluation itself is conducted in a conflict-sensitive way. Conflict analysis, as the basis for evaluative analysis, is a key aspect of conflict prevention and peacebuilding evaluations, regardless of the design and methods used. It is as important for evaluations using randomised control groups, regression analyses, surveys, and large sample sizes, as it is for qualitative evaluations with in-depth case studies and focus groups. (The use of conflict analysis is covered in more detail in Chapter 3 and Annex A).

Conflict sensitivity

Conflict sensitivity refers to the ability of an organisation to *a)* understand the context in which it is operating, *b)* understand the interaction between the intervention and that context, and *c)* act upon that understanding in order to avoid negative impacts and maximise positive impacts on the conflict (CDA, 2009). All activities in a fragile and conflict-affected setting must be conflict sensitive. The principles of conflict sensitivity, adopted by the OECD in 2001, assert that international assistance must, at a minimum, avoid negative effects on conflict – "do no harm" – and, where possible, make a positive contribution to conflict prevention and peacebuilding. Fragile States Principle 2 (OECD, 2007) reiterates the commitment to conflict sensitivity, emphasising the importance of basing interventions on strong conflict and governance analysis in order to avoid inadvertently aggravating social tensions or exacerbating conflict.

Conflict prevention and peacebuilding policies, projects and programmes, and development or humanitarian activities in conflict settings sometimes do cause harm, often unwittingly as in the example given Chapter 1 (Box 1.3). When assistance does cause harm in a situation of conflict and fragility, it produces direct or indirect effects that aggravate grievances, increase tension and vulnerabilities, and/or perpetuate conflict and fragility in some way. Such effects may be the result of a project or programme engagement – i.e. how its humanitarian or development outcomes contribute to peace or affect conflict. However, they may also spring from the operational aspects of an engagement (Uvin, 1999ab) – i.e. how, where, and when donors and agencies operate and how they implement and distribute aid.

As a policy or programme should be conflict sensitive, so should the evaluation process itself. Evaluations carried out before, during, or after a violent conflict must be conflict sensitive because they are themselves interventions that may impact on the conflict. In this respect, it is important to understand that questions asked as part of an evaluation may shape people's perception of a conflict. Evaluators should be aware that questions can be posed in ways that reinforce distrust and hostility towards the "other side". Evaluators should keep in mind that the way they act, including both the explicit and implicit messages they transmit, may affect the degree of risk.

Moreover, the evaluation process itself may actually put people in danger. A number of the evaluators who contributed to this guidance spoke of incidents where someone they

had questioned in the course of their evaluation work had been arrested or otherwise threatened. Measures should be taken to avoid this. For example, in one evaluation the evaluation team leader decided that the names of its local members should not published in the report, because of possible repercussions they could face as a result. Their identities were protected and the local experts operated instead as external resource staff and key informants. In some cases, such as the real-time evaluation of Denmark's humanitarian aid in south-eastern Somalia (Polastro *et al.,* 2011), it may be considered more prudent and effective to rely on local staff or national teams that can more easily travel in dangerous zones – though their safety must also be protected.

It is especially important to consider the safety of interpreters and other local staff, partners and beneficiaries, whom evaluators may inadvertently expose to greater risks than they themselves face. International evaluators leave after a short while, which may influence the risks they are prepared to take. Local people stay, however, and face possible reprisals. Such risks should be identified and addressed at the outset of the process and included in the planning and implementation of the evaluation. Doing so is the responsibility of evaluation commissioners and team leaders and a requirement of conflict-sensitive, ethical evaluation. Evaluators and commissioners should discuss and take appropriate measures to ensure conflict sensitivity, the ethical conduct of the evaluation and the protection of those involved. A thorough, up-to-date understanding of the conflict is the first step in a conflict-sensitive evaluation process. The evaluation report must explain what measures were or were not taken to ensure the conflict sensitivity of the evaluation itself and any impact that taking or not taking them may have had on the results of the evaluation.

Evaluating conflict sensitivity (and effectiveness)

It is important to understand that conflict sensitivity does not automatically deliver an effective peace programme or policy. A conflict-sensitive intervention is not necessarily effective in addressing drivers of conflict and fragility. Nor are explicit peacebuilding interventions necessarily conflict sensitive. For example, a reconstruction programme that rebuilds destroyed homes and provides small income-generation grants to returning refugees and internally displaced persons may avoid "doing harm" and try to rebuild relationships across conflict lines. It sponsors inter-ethnic dialogue between returnees and host community members, provides "balancing grants" to the host communities for priority community infrastructure or income-generation projects, and sponsors sports and cultural events for youth. It succeeds in ensuring that aid does not disproportionately benefit one group, and supports rebuilding of relationships among some community members. However, while it may be conflict sensitive, the reconstruction programme may not be effective peacebuilding as such, insofar as its activities do not address the drivers of the conflict, which could be, for example, impunity and injustice or conflicting visions of the future.

In assessing conflict sensitivity, it is important to look at the extent to which the intervention aggravates or mitigates grievances, vulnerabilities or tensions. For interventions that do not have explicit peacebuilding goals, evaluators would assess the effects of the development or humanitarian outputs and outcomes (e.g. infrastructure development, a more operational police or judicial system, etc.) on the drivers of conflict or fragility. For example, a poverty reduction programme may have positive development results, but a thorough conflict analysis might reveal that, while the programme reduced levels of poverty overall, one group gained more than another, causing deeper resentment

among excluded groups. If poverty reduction strategies helped achieve greater equity, they might contribute to peace.

In addition, all activities, whether explicitly aimed at peacebuilding or not, should be examined to assess their conflict sensitivity. One of the more widely used conflict sensitivity tools, the *Do No Harm Framework* (Anderson, 1999), draws attention to the unintended consequences of aid planning and practice. Although it was originally developed for humanitarian aid it is also regularly applied to development and peacebuilding interventions. It identifies five ways in which operational components of an intervention may affect a conflict:

- theft/diversion: fuelling the conflict with stolen or diverted goods/funds;
- market effects: changing local markets with an influx of outside goods;
- distribution: distributing goods along the lines of the conflict;
- substitution effects: replacing existing functioning systems or structures;
- legitimisation: giving legitimacy to a group or leader by working with them.

It also identifies four ways in which the behaviour of agencies, especially those implementing programmes, sends messages that reinforce the modes of warfare or, alternatively, non-conflictual relations. These include behaviour that:

- conveys respect or disrespect to people and communities,
- communicates an agency's willingness or unwillingness to be held accountable,
- treats people in ways that are perceived as fair or unfair,
- demonstrates transparency or lack of transparency.

Negative patterns can undermine an organisation's efforts and put its staff in danger, lead to relationships that are antagonistic and untrusting, and make partners and communities feel humiliated. In extreme cases, violating the principles of respect, accountability, fairness and transparency can lead to violence against an organisation or within the community.

Evaluators may need to examine the target agency's own ways of working to determine whether the intervention is conflict sensitive. This would include examining inadvertent impacts of decisions about staffing, criteria for selection of beneficiaries, selection of local partners, relations with local authorities (including military actors and government), and processes and procedures for distributing aid (*ibid.*). Often, simple decisions about hiring – such as requirements regarding language – can result in staff that is disproportionately drawn from one conflict group. Similarly, seemingly objective criteria for the selection of beneficiaries (e.g. needs) can result in one group obtaining much more assistance than another and, consequently, contribute to escalating tensions. While the implication is not that donors or implementing staff abandon their criteria or redistribute aid, they must be aware of unintended conflict effects and develop options within the programme to mitigate them (*ibid.*).

Being conflict sensitive and evaluating conflict sensitivity are two imperative dimensions of evaluating conflict prevention and peacebuilding work. A clear, critical assessment of an activity or a policy's impacts will cover both intended and unintended consequences and thus offer insights into the sensitivity of the activity under evaluation. Evaluators can help assess whether or not the standard of conflict sensitivity has been achieved – as well as provide insights on how to improve sensitivity.

© OECD 2012

Chapter 3

Preparing an evaluation in situations of conflict and fragility

This chapter considers how evaluation commissioning agencies and planners may set up an evaluation. Its base premise is that effective preparation makes for effective evaluation. It thus examines each of the key preparatory steps, looking first at how to define the purpose of an evaluation and how to conduct (or commission) a conflict analysis. The chapter then goes onto discuss how to identify the key questions an evaluation must ask. It examines timing and logistics, co-ordination with other actors, selecting evaluation criteria, management, methods, the evaluation team, and the dissemination of evaluation results.

This chapter is designed to be of interest to those who commission evaluations (programme managers, staff or evaluation offices) and evaluation managers, as well as staff and decision makers involved in selecting evaluation topics and questions. It sets out some key steps in preparing an evaluation. The first three steps – purpose, conflict analysis, and scope – may usefully be considered together.

Summary of key steps for preparing an evaluation	
✓ **Define the purpose of the evaluation**	Develop terms of reference
✓ **Analyse the conflict context**	
✓ Consider gender equality	
✓ **Determine the scope of the evaluation**	
✓ **Decide on** evaluation criteria	
✓ Outline key evaluation questions	
✓ Select evaluation approach and method to fulfil purpose	
✓ Take timing and logistical issues into consideration	
✓ Co-ordinate with other actors	
✓ Determine how the evaluation will be managed	
✓ Select and contract the evaluation team	
✓ Prepare to disseminate evaluation results	
✓ Control quality	

Defining the purpose

Every evaluation, regardless of the context, should begin with the question: What is this evaluation meant to ascertain and how will this information be used? Defining the purpose and objectives of an evaluation is the most important planning step. If the purpose is not clear, the evaluation will not be. Commissioners should think about: Who is the intended audience? Who is to receive the findings and what will they do with the results? What kind of information is needed?

Evaluations may have a number of different, sometimes concurrent, purposes. Accountability and learning are two of most frequent, and most evaluations combine them.

- **Accountability** seeks to find out whether an activity has been performed as intended and/or whether it has achieved the expected results.

- **Learning** looks to provide evidence and improve knowledge of results and performance, which can help improve ongoing or future activities and increase understanding of what works, what does not, and why.

Box 3.1 gives some examples of the accountability and learning objectives of evaluations, as described in evaluation reports.

Box 3.1. **Defining the purpose of an evaluation**

- **Accountability**: "The purpose of the evaluation is to assess whether Norway has, with its transitional assistance, contributed to increased security (and stability) in Haiti, and whether gains achieved are likely to be sustained." (Norad, 2009).
- **Accountability:** "This report was prepared to ascertain whether Asian Development Bank policy conditions had been met and whether they led to achievement of the Tajikistan Post Conflict Infrastructure Programme's stated objectives or purpose." (ADB, 2007).
- **Learning:** "The overall objectives of this project were, first, to develop a method for assessing the impact of development co-operation in conflict zones, and second, to apply this method in North East Afghanistan." (Böhnke *et al*, 2010).
- **Accountability and learning**: "The joint evaluation of conflict prevention and peace building in the Democratic Republic of Congo has a double purpose: to provide accountability to the public and to decision makers in development co-operation, and to generate lessons for improvement. The emphasis is on the learning side, with a view to developing more strategic policies and programmes." (Brusset *et al,* 2011).

Source: Development Evaluation Resource Centre (DEReC) website, *www.oecd.org/dac/evaluationnetwork/derec.*

The *DAC Principles for Evaluation of Development Assistance* (OECD, 1991) state that "to have an impact on decision-making evaluation findings must be perceived as relevant and useful". Evaluations involve real costs, including the use of resources which could otherwise be deployed elsewhere, and should therefore be judged on the value of the information they provide. Usefulness is an important principle in evaluation. Use can take many different forms, before and during the implementation of an evaluation, or even many years after. In some cases decision makers use the findings to change or modify a programme directly, based on the recommendations presented. But in many cases use is less direct. An evaluation may contribute, along with other evaluations and research, to building up general knowledge over time on a particular topic, for instance. Behavioural or organisational changes may be caused by engaging in the evaluation process itself. Factors that may influence use of evaluations, and can be kept in mind when planning an evaluation, are the institutional environment (incentives and capacity for use), the relevance of the evaluation (timing, involvement of stakeholders, credibility), and the quality of dissemination (evaluation product, communication channels and mechanisms) (Feinstein, 2002).

For evaluations involving multiple stakeholders, a shared understanding of the overall goal is crucial. In order to ensure such shared understanding and, later, the usefulness of the evaluation, involving stakeholders as part of the preparation and planning process is recommended. However, their degree of involvement depends on the evaluation's design and purpose. An evaluation focused on learning is likely to be more participatory, whereas stakeholders would be less involved in an evaluation with an accountability purpose.

Analysing conflict

Conflict analysis helps identify the causes, drivers and dynamics of conflict and fragility. It provides an analytical framework for understanding the complex, changing context in which an intervention is implemented. A conflict analysis identifies the key factors relating to conflict and fragility and the linkages between them, pointing to the sources and dynamics of violence as well as peace. A good analysis of conflict and fragility should include (or be linked to) an in-depth analysis of the political economy and broader development context (Kennedy-Chouane, 2011).

Evaluators will always need to have an understanding of the conflict as the basis for their work, though they may not necessarily need to perform a conflict analysis themselves. The evaluation could be based on analysis provided by the evaluation target itself, the commissioning agency, an independent research institute or consultant, or a participatory process with stakeholders. An analysis may also be performed by the evaluation team as an early step in the evaluation process or used by commissioners to refine the evaluation scope and define key questions.

Box 3.2. **Two examples of the use of conflict analysis – Democratic Republic of Congo and Sri Lanka**

In the Democratic Republic of Congo, the conflict analysis – a combination of scientific research and workshops in the field – identified four important drivers of conflict during the inception phase. The four drivers – land ownership, weakness of the state, security sector, natural resources – helped delineate the scope of the evaluation. The drivers also helped evaluate the relevance of interventions (Did they target the right drivers of conflict?) and their impact (Did the conflict prevention and peacebuilding assistance have an effect on these drivers?) (Brusset *et al.*, 2011).

In Sri Lanka, the study used the existing comprehensive strategic conflict assessments conducted in 2001 and 2005 as a point of reference and background for the analysis. Though the team recognised that the conflict had deepened since 2005, it felt its root causes remained unchanged and were sufficiently covered by the earlier analyses (Chapman *et al.*, 2009).

Source: Brusset *et al.* (2011) and Chapman *et al.* (2009).

As described in Chapter 1, a thorough understanding of the context of conflict and fragility should be part of the design and management of all interventions. If a conflict analysis has been carried out as part of developing a donor's strategic engagement or programme design, the evaluation team will need to review the analysis and assess its quality and relevance at the outset of the programme and how it was adapted (or not) over time. Evaluators will need to consider whether the underlying analysis (explicit or implicit) was sufficient and accurate, whether it was effectively translated into relevant strategies and objectives, and whether it was adapted to the situation of conflict and fragility over time. Tips for reviewing a conflict analysis are provided in Box 3.3.

There are many different models, tools and frameworks for conflict analysis used by development donors and others working in and on conflict and fragility. The aim is to gain a broad and deep understanding of the context in order to evaluate the intervention in

question. Notwithstanding the diversity of different models of conflict analysis, there is growing consensus on what is required in a good analysis:

- It distinguishes between the structural causes of conflict and fragility (both issues and people) and dynamic events and trends.
- It identifies positive as well as negative forces affecting the conflict (and chances for peace).
- It prioritises drivers of conflict and fragility and identifies which ones can be influenced by external action.
- It is operationally useful, and reflected in programme design, monitoring and evaluation.

Nevertheless, in order to avoid the trap of becoming too comprehensive (and thus difficult to operationalise), it is important to distinguish those elements of the broader context that directly influence the conflict and how they do so. For instance, if poverty is identified as an important factor in the context, conflict analysis should identify which aspects of poverty influence tensions, resentments, and violence and in what way. Box 3.3 outlines more key elements to look for in a conflict analysis.

Drawing on the outcomes of the conflict analysis, evaluators can assess the relevance and impacts of the activity or policy in question. For instance, the outcomes of the analysis will help to gauge whether or not an intervention addresses the relevant needs of the context, i.e. the causes of conflict and fragility. Additional information will, however, be needed to evaluate all dimensions of relevance and impact. For instance, the relevance of donor activities to overall country strategies or donor priorities may not be revealed through a standard conflict analysis and will have to be captured with other data, including programme documents and information on policies.

One way of developing an analysis is to involve a range of stakeholders early on in the evaluation process. As it is not always possible to obtain all the competing perspectives from the different parties at the same time, it may be advisable to interview people separately to gain a deeper, wider understanding of the situation. However, evaluators should be aware that it will likely be difficult to gain consensus on the nature of the conflict as contending groups will not agree. This, of course, is a natural characteristic of conflict – and competing interpretations of history and causes may be an important dimension that the analysis captures. The outcomes of interviews, therefore, should be triangulated with secondary sources such as policy documents, programme/project notes, and grey literature such as reports from research institutes and think tanks. In interviews it can be particularly useful to engage in discussion to prioritise drivers of conflict and fragility, to understand which drivers are really important (and which less so), and to make general assertions more specific.

Most evaluations look at an intervention or overall engagement in a single country or conflict region. In that case, the conflict analysis focuses on understanding that particular conflict, and also examines sub-regional or local conflict dynamics as relevant. For evaluations that involve analysis of activities across various conflict contexts – e.g. a thematic evaluation looking at women's role in peace processes or an evaluation of a disarmament programme that operates in several different post-conflict countries – the analysis of conflict will be approached differently. Analysis could draw on existing research and empirical evidence about the (assumed) connections between the type of activities and violent conflict or state fragility in general. Case studies and comparative analysis can be

Box 3.3. **Checklist for reviewing a conflict analysis**

If a conflict analysis has been carried out as part of developing a donor's strategic engagement or programme design, the team will need to review the analysis and assess its quality and relevance at the outset of the programme and how it was adapted (or not) over time. In this process, the evaluation team should pose the following questions:

1. Given the resources and capacities of the agency or organisation being evaluated, was the appropriate conflict analysis approach or tool chosen to guide the design and implementation of the programme(s) or policy(ies)? Did the analysis generate adequate information for determining the relevance of the intervention to the needs of the peacebuilding process; to the effectiveness of the programme designs and implementation; and to an assessment of the appropriateness of the theory of change?
2. Was the analysis kept up-to-date from the time the programme or policy was designed through the period of time under evaluation? Does it capture the evolution of the conflict in a way that can be used to look at relevance and longer term impacts? (If not, the evaluation team may need to update the analysis.)
3. Was the process of conflict analysis appropriate and effective?
   a) Was the analysis conducted by skilled people with an understanding of conflict and of the?
   b) Did the analysis gather information from a wide range of sources? Did it include perspectives from all the main stakeholders in the conflict?
   c) Was the analysis conducted in a conflict-sensitive manner? For example, did it ask questions in a way that avoided exacerbating divisions? If the analysis was conducted by convening stakeholder workshops, did the facilitators possess, or lack, sufficient skills to engage conflicting parties in productive discussion? Did the analysis process put researchers (and local partners) at risk by sending them into insecure areas? Did it put interviewees at risk by exposing them to retaliation?
4. Was the analysis done at the appropriate level? For example, if a programme was to be initiated at the provincial level, was a national analysis supplemented by an analysis of conflict dynamics within the province?
5. Were the conclusions reasonable? Were critical elements missing from the analysis? To what degree was the analysis shaped by the expertise of the agency or their general beliefs about how to bring about positive change?

Was the analysis linked to strategy? Did it actually inform implementation and activities?

used to test these hypotheses and assess the intervention's relevance, effectiveness and impact, as discussed in Chapter 4.

Given the many different conflict analysis models and frameworks, evaluation commissioners and managers need to give conflict analysis careful thought. The analysis method selected should be well adapted to the context, the evaluation scope, and available resources. It follows that funders of evaluations should ensure that resources for the conflict analysis are proportional to the task envisaged and that the evaluation team has the necessary skills to analyse conflict.

Broad questions that can be included are listed in Table 3.1, while Annex A provides further discussion on different approaches and links to conflict analysis resources.

Table 3.1. **Some key questions for conflict analysis**

Profile	What is the political, economic, and socio-cultural context? What are the emergent political, economic and social issues? Are there important regional/international dynamics? What are the geographic dimensions? What areas that are prone to conflict and fragility, or affected by them, can be situated within the context? Is there a history of conflict?
Conflict causes and potentials for peace	What are the structural causes of conflict and fragility? What issues can be considered as proximate or dynamic causes of conflict and fragility? What triggers could contribute to the outbreak or further escalation of violence? What are the strategies or habits for dealing with conflict that contribute to violence? What new or emerging factors contribute to prolonging conflict and fragility dynamics? Have original causes shifted due to events during war and mass violence? What factors can contribute to peace and stability? What existing factors bring people together and can be built upon or reinforced? What are the most important drivers of conflict and peace? Which factors have the greatest influence on the situation?
Actors	Who are the main actors (people who perpetuate or mitigate the situation of conflict and fragility)? How do they contribute to or mitigate conflict? What are their interests, goals, positions, capacities and relationships? What capacities for peace and stability can be identified? Who can make a difference? What actors can be identified as "spoilers" (those who benefit from ongoing violence or who resist movement towards peace and stability)? Why? Are they inadvertent or intentional spoilers?
Dynamics and future trends	What are the relationships and dynamics among the key drivers of conflict and peace? What are the current conflict and fragility trends? What are the negative reinforcing cycles? What are the windows of opportunity? What scenarios can be developed from the analysis of the conflict and fragility profile, drivers and actors? How might different scenarios play out given likely future developments (in the short and long run)?

Source: Adapted from International Alert (2007a) and Paffenholz and Reychler (2007).

Deciding the scope of the evaluation

The scope of the evaluation should be clearly defined. The scope specifies the issues, funds, or types of interventions to be covered, including the time period and geographical coverage. When determining an evaluation's scope, it is important to clarify and agree which types of aid it will cover and how. Evaluations may consider all or part of aid in a particular context, including explicit peacebuilding efforts, other forms of development assistance and humanitarian aid. Table 3.2 provides an overview of the hierarchy of evaluation scopes and some real-life examples for reference.

Table 3.2. **Examples of evaluation scopes**

Type of evaluation	Definition	Example
System-wide or country-level	Evaluation of the response by all (or most) international partners in a particular country or to a particular armed conflict or outbreak of violence.	Multi-donor evaluation of Support to Conflict Prevention and Peacebuilding Activities in Southern Sudan 2005-2010 (Bennett *et al.*, 2010).
Partial system	Evaluation of a part of a system (e.g. thematic or sector study), which may include cross-country or cross-conflict analysis.	Joint Evaluation Programme on theme of Support to Displaced Persons (Borton et al., 2005). Evaluation of the German Civil Peace Service (Paffenholz , 2011).
Single-agency response	Evaluation of the overall response to a particular country or armed conflict by one international partner (funding, channelling, or implementing agency).	Evaluation of Norwegian Support to Peacebuilding in Haiti 1998-2008 (Norad, 2009).
Single project	Evaluation of a single project, programme or policy undertaken by a single agency.	Evaluation of the "Open Fun Football" School Programme in the Balkans region (Danida, 2011).

Source: Examples drawn from the DAC Evaluation Resource Centre (DEReC), *www.oecd.org/dac/evaluationnetwork/derec*.

The scope should also spell out specific policies to be addressed in the evaluation – country memorandums of understanding or (joint) donor engagement strategies, for example. Conflict analysis can inform the process of determining the scope of the evaluation.

Box 3.4. **Using conflict analysis to inform the scope of an evaluation in the Democratic Republic of Congo**

A multi-donor evaluation was launched in 2008 to assess the role of external partners in supporting peacebuilding and conflict prevention in the eastern Democratic Republic of Congo. The use of a conflict analysis was very helpful for determining what key conflict factors – both the obvious and the less obvious ones – should be covered in the evaluation. At first, the evaluation focused on sexual and gender-based violence, child soldiers, and natural resources, which the commissioning evaluation departments considered to be key factors in the conflict. At that time (2008), these were generally accepted as important but the choice of those three factors was not based on a conflict analysis. Once the evaluation got underway, the team used conflict analysis to identify land issues and the weakness of the state as major conflict drivers, and these became part of the evaluation scope.

Source: Brusset *et al.* (2011).

To tailor an evaluation's scope to its purpose and available resources, planners should ask these questions: What activities and policies will be covered? How far along the "results chain" will the evaluation go? Will it look for immediate impacts or on broad conflict dynamics? The answers to these questions will influence the selection of criteria and methods as described in the next two sections.

Selecting evaluation criteria

When planning an evaluation and drawing up its terms of reference, commissioners will need to determine which criteria will be analysed. The criteria to be examined are usually included in the key evaluation questions (see below) and will make up the main analytical content of the evaluation.

The five OECD DAC criteria for evaluating development assistance – relevance, effectiveness, efficiency, sustainability and impact – are usually considered, though it may be more manageable to focus the evaluation on looking at a few criteria in-depth, depending on the evaluation purpose and intended use. Additional considerations that may be particularly relevant to situations of conflict and fragility, namely coherence and co-ordination, could also be subject to examination. Chapter 4 features a section, "Criteria for evaluating interventions", that discusses use of the criteria in an evaluation.

Outlining key evaluation questions

Evaluation managers should develop a list of questions (or lines of inquiry) that an evaluation will answer. The type of intervention, the stage of implementation, and what the evaluation hopes to achieve determine the specific evaluation questions. In some cases, questions will be specific at the outset of an evaluation. In others, general questions will be refined through an iterative process during the evaluation. When considering

evaluation questions, evaluation managers should also think about which methods to apply for answering the questions and whether those methods are feasible in the available time and budget?

When evaluating peacebuilding or statebuilding support and development interventions in fragile or conflict-prone contexts, evaluators might (in addition to assessing conflict sensitivity) pursue the following lines of inquiry:

- Is the intervention addressing the driving factors of the conflict? Does (or could) it address key tensions that have been identified as key factors in past, current, or possible future conflict?
- Has an analysis of conflict and fragility dynamics been undertaken and has it influenced programming and implementation choices?

Humanitarian activities in conflict situations are guided by the core principles of neutrality and impartiality, and other *Principles and Good Practices of Good Humanitarian Donorship* agreed in Sweden in 2003 (GHD, 2003). Evaluations are likely to focus on assessing the extent to which a humanitarian intervention abided by the principles and the results it produced. And because humanitarian actions may have unintended (positive or negative) influence on conflict dynamics, evaluations must also consider conflict sensitivity. If the scope of an evaluation takes in all external engagement in a particular country, the relevance and effectiveness of humanitarian aid in relation to the conflict and fragility dynamics should also be considered. When examining the entire portfolio of assistance in a country affected by conflict, fragility, or prolonged humanitarian crises, evaluators might include humanitarian interventions in their analyses in order to assess the overall impacts on peace and conflict of interventions by external partners. They might also wish to consider the balance between humanitarian aid and other types of assistance.

Some questions an evaluation might ask about humanitarian aid include:

- Does the intervention avoid creating tensions within the crisis-affected community; between displaced people and host communities; between agencies over the type and quantity of assistance?
- Does the provision of humanitarian aid impact the role and legitimacy of the state or have an influence on statebuilding processes?
- Is there coherence between humanitarian activities and other types of assistance?

Because peacebuilding and statebuilding interventions may affect men and women differently, commissioners will often include questions on gender inequalities. Questions may focus on disparities in the family, the community, the marketplace, the state, and consider issues like gender-determined division of labour and role assignment; unequal access to and control over resources, benefits, and services; disparate participation in public and private spaces; and gender-specific practical or strategic needs such as protection from violence. A gender analysis can form the basis for studying gender dimensions in the evaluation.

It is important to be realistic about what an evaluation can achieve, particularly when selecting evaluation questions. The testing phase of this guidance revealed a tendency for commissioning agencies and consultants to have overly ambitious expectations in respect to scope, content and timelines. For example, some evaluation terms of references contained dozens of evaluation questions, making it difficult for the evaluation team to answer all of the questions properly within the expected timeframe. One solution may be to provide a few broad evaluation questions at the outset, which can be examined in

greater depth, or added to, as part of the evaluation analysis. In an evaluation of the German Civil Peace Service (Paffenholz, 2011), the evaluation team first did a pilot of one of the eight case studies (Uganda) and used that experience to fine-tune the evaluation questions. They realised that several questions in the terms of reference could not be answered in the individual case studies and had to find other ways to address those questions.

Incorporating gender equality and women's empowerment

Those planning an evaluation will need to determine how it will cover gender issues. Field experience and extensive research show that women and men and boys and girls experience, engage in, and are affected by violent conflict in different ways. Educational levels, responsibilities and mobility are some of the factors that can vary with gender and may affect the resources available to women and men in a conflict or post-conflict situation. Conflict itself can play a role in forming or changing a society's understanding of gender roles – i.e. what it expects of different individuals in a given context. In many cases, conflict increases the burden placed on women. Systematic violations of women's rights and their exclusion from economic, social and political spheres are barriers to development and may affect conflict dynamics. A clear, critical understanding of gender equality within a particular conflict context is, therefore, important for policy makers and practitioners, as well as for evaluators.

United Nations Security Council Resolution 1325 (UNSC, 2000) was the first Security Council Resolution to link women's experiences of conflict to the international and peace security agenda. It established that equality between men and women was essential to achieving and sustaining peace, and that equal participation in peace processes and post-conflict reconstruction efforts was critical to peacebuilding and statebuilding. It also called for the protection of women and girls and the inclusion of gender equality considerations in peacekeeping operations and training.

In line with Resolution 1325, gender dynamics are part of conflict analysis. This requires differentiating between the roles, experiences and perspectives of men, women, girls and boys, as well as among women and men of different social, ethnic, religious or economic groups. Such analysis should not, however, fall into the trap of gender-based stereotypes. In the past, development agencies viewed women and girls primarily through the lens of victimhood. As a result they were too often left out of peacebuilding and statebuilding processes. In light of UNSCR 1325, there is growing consensus on two dimensions critical to understanding women's roles in conflict situations: the targeting and victimisation of women and girls and the key part they play in peacemaking and rebuilding societies. This approach can help to ensure that peacebuilding processes take women's needs into account and include making space for women in government and in key post-conflict decision-making processes. That being said, it should be acknowledged that women may also be perpetrators of violence, just as men may be victims.

Those commissioning an evaluation should determine how it addresses gender-related issues in accordance with its focus and the activity or policy it is assessing. Planners may choose to make gender a cross-cutting theme or specific focus (some development agencies have particular requirements for the coverage of gender equality in evaluations). A programme that does not have adequate understanding of different gender needs and roles, or fails to adjust to them, may lack effectiveness, impact and relevance. Such issues might be included in the evaluation questions. "Criteria for evaluating interventions" in Chapter 4 gives several examples of questions on gender equality.

Evaluation managers might decide to include certain requirements on gender equality and women's empowerment in the terms of reference as a way of ensuring that they are integrated into the evaluation's objectives. For instance, commissioners may request that the evaluation team include a gender expert and use gender-disaggregated data and gender-sensitive indicators.

An example of how gender could be incorporated into a conflict evaluation would be the evaluation of an infrastructure reconstruction project in a conflict setting that also looks at whether jobs created by the project have affected the livelihoods of working age men and women differently and whether this impacted on conflict drivers. As part of their efforts to ensure a conflict-sensitive approach, the evaluation team should examine how considerations of gender equality may affect their own work – in the make-up of their team or their engagement with stakeholders, for example.

Looking at the big picture

Experience in fragile and conflict-affected contexts has resulted in a growing emphasis on the need to look beyond individual projects, development assistance, and individual actors in order to understand peace and development processes more broadly. An individual activity may successfully achieve its short-term outcomes, such as training police officers or providing new livelihoods for former soldiers. However, these micro-level successes have widely been seen as not "adding up" to real progress towards peace. There is often a paradox between programme reports or self-evaluations that show successful programmes with good results and a simultaneous lack of progress towards peace – or even escalation of violence – at the macro level (sometimes called "peace write large").

The success of individual development or humanitarian activities and the outcome of overall peace and statebuilding processes generally depend not on the actions or strategies of a single funding or implementing agency, but on other factors. For instance, external shocks, government policy decisions or the diplomatic pressure that the international community exerts – or does not exert – on governments and warring parties might drive conflict and peace dynamics. Such factors are beyond the scope of the evaluated activity, but evaluators must nonetheless consider them in order to draw reasonable conclusions and attribute results. Evaluations that focus narrowly on donor engagement or look only at the programme or sector level may fail to identify important system-wide effects or constraints.

It would not, of course, be realistic for every single evaluation to cover the entire policy arena or all dimensions of conflict and fragility in a given context. Activity or programme-level evaluations are also valuable. Nevertheless, development agencies should also plan evaluations that capture strategic issues or may ask evaluation teams to examine questions related to the broader context during an evaluation of a single programme or activity.

Selecting the best-fit evaluation methodology

There is no single blueprint methodology for evaluating donor engagement in fragile and conflict-affected situations. Rather, the golden rule is to apply the right tools and methods to the right questions. Methods should be chosen according to the evaluation purpose and key objectives and should involve a credible approach to attribution that avoids potential biases. The complex nature of interventions in fragile and conflict-affected

situations generally makes it necessary to combine different methodologies in order to answer the evaluation questions. Many favour a mixed-method approach, using both qualitative and quantitative methods and data. All the evaluations that tested this guidance used such mixed-method approaches. Planners should examine questions of methodology, deciding, for example, where there will be comparison groups (to inform discussion of counterfactuals), single case studies, and time or single series data.

Extensive literature exists on the strengths and weaknesses of evaluation methods and their applicability to different contexts and purposes. Several types are listed in the glossary. An overview of methods and their suitability to conflict prevention and peacebuilding work may be found in Church and Rogers (2006) and in OECD (2008b).

Encouraging participation of both women and men and knowing the informal rules of communication between men and women, is central to selecting a gender-sensitive approach. The incorporation of both women and men in the sample or study population should be ensured and potential obstacles to women's participation in the evaluation addressed. For instance, it could be difficult for evaluators to speak directly with women and women may not express themselves freely in the presence of men. The methodological implications of these gender dynamics should be considered.

Dealing with timing and logistics

Schedules and evaluation plans are often decided well in advance. However, the timing for evaluating conflict prevention and peacebuilding interventions should be determined not only by the phase of the policy, programme, or project cycle, but also in relation to current conflict realities. The timing of the evaluation should be appropriate to the current dynamics of conflict and fragility and useful for informing policy discussion and/or programme adjustments (according to objectives). Commissioning organisations may have to adjust their expectations in the light of conflict and fragility-related constraints.

The terms of reference should be clear about realistic time frames. To identify the right time and good entry points for an evaluation, the questions below should be considered, bearing in mind the outcomes of the conflict analysis. Clear terms of reference will help ensure the conflict sensitivity of the evaluation process itself, particularly how it proposes managing logistics and timing.

- 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.).
- Would an evaluation at this moment be disruptive to the policy, project, or programme itself?
- 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?
- 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, programme or project and the staff and stakeholders.
- Has the activity been in place long enough to provide useful experience and learning? Is the assessment of outputs, outcomes, and impacts based on a realistic time frame?

- How long has it been since any previous evaluation or review was performed (either of a donor's own activities or of relevant or similar activities of other donors)?
- Are there any logistical issues that must be taken into consideration (security restrictions, election process, weather patterns, major national holidays, access to transport, etc.)?

Box 3.5. **Conflict-related constraints on evaluations in Sri Lanka and the Democratic Republic of Congo**

In Norway's evaluation of peace efforts in Sri Lanka during the period 1997-2009, the team faced a challenging task. The evaluators were granted full access to the Norwegian Ministry of Foreign Affairs' archives and Norwegian individuals involved in the peace process. However, they could not gain access to a number of key people in Sri Lanka. They included senior figures in the Liberation Tigers of Tamil Eelam (dead), second-level cadres (in prison), and the incumbent government. The team sought to compensate by studying secondary sources, such as published research (including the team's own), unpublished reports, and media coverage. In addition, international and national actors, experts and observers were interviewed.

In the Democratic Republic of Congo, a first attempt in 2008 at evaluating peacebuilding support in Eastern DRC was cancelled due to a resurgence of violence. A rebel group, the National Congress for the Defence of the People, launched an offensive in North Kivu and the safety of the evaluators in the field could not be guaranteed. In 2009, there was a fresh attempt at conducting an evaluation. This time the set-up was different. Greater emphasis was placed on policy analysis and document-based study and interviews. Had there been a fresh upsurge in violence, the new set-up would have made it possible to continue the evaluation. However, without the field missions the evaluation would have also shifted in focus, looking at donor policy only and not assessing results in the country.

Source: Norad (2011) and Brusset *et al.* (2011).

Co-ordinating with other actors

In line with the evaluation quality standards drawn up by DAC (OECD, 2010c) and the widely endorsed principles of aid effectiveness (OECD, 2005), evaluation work should be co-ordinated where possible and purposeful. To facilitate co-ordination, it is important to examine the institutional, organisational, and project-level context of the intervention in order to identify key stakeholders. Actors may include development agencies, bilateral donors, and multilateral institutions, providers of South-South assistance, implementing agencies, non-state actors, civil society, humanitarian actors and military forces. Commissioners should consider the interests different actors might have and contributions they could make in terms of data and decide if and how they might be involved in the evaluation. For example, non-governmental organisations that play a major role in the sector could be invited to serve on an evaluation reference group to lend the evaluation additional scope and reach. Or military actors that have access to data about the security situation may be asked to contribute this information without actually joining the evaluation process. For an evaluation looking at support to education, the Ministry of Education could be engaged and contribute to defining key evaluation questions and understanding potential links between education and conflict.

Increasingly, development agencies, humanitarian organisations and security forces are working together in situations of conflict and fragility. The current emphasis among

many countries on "whole-of-government" approaches can often involve a great variety of actors from diverse backgrounds in any one evaluation. Co-ordinating such a diverse cross-section requires special consideration when setting up the management structure. If handled carefully, bringing together different development assistance players in a single evaluation can be a learning experience in that it broadens the scope of analysis and affords an opportunity for assessing differences in intervention methods and theories of change. However, a very large number of actors, ill-defined roles, or unclear objectives can make co-ordination a real problem.

"Criteria for evaluating interventions" in Chapter 4, which discusses the importance of ensuring the coherence and co-ordination of an evaluated activity, also considers the involvement of different actors.

Working with local and country stakeholders

Evaluation managers may decide to invite local stakeholders (country offices or embassy staff, national governments, civil society organisations, beneficiaries, implementing partners) to take part in planning and conducting an evaluation – especially when learning and using the results are of the essence. It is generally accepted that there is a need for external partners to increase the involvement of local people and intended beneficiaries in evaluation. Local involvement may contribute to ensuring a more transparent, stronger relationship between external actors and local communities, in line with the *Fragile States Principles* and extensively borne out by evaluation experience. Engaging with knowledgeable local people and those targeted by programmes can provide critical input for understanding the context and conflict and carrying out the evaluation analysis. Involvement of people from different sides or with different perspectives on the conflict can be critical to understanding links between the intervention and conflict dynamics. Nevertheless, planners must take great care when deciding whom to involve, and how to involve them, in the context of fragility and violent conflict. The need to protect those involved and safeguard the objectivity and impartiality of the evaluation may influence such decisions.

Development interventions that affect daily activities, resources, roles and responsibilities, opportunities and rights of the beneficiaries may have different implications (intended and unintended, positive and negative) for women and men and boys and girls. Evaluation managers must make sure that different points of view are included throughout the evaluation process. This may require considerations about what to analyse, which questions to ask stakeholders, what data collection procedures to adopt, what type of report to write and how to disseminate it. In some contexts it requires including female members in the evaluation team.

Evaluation planners will also need to determine how to handle the participation of partner country government institutions and what level of involvement is appropriate and useful in a specific conflict context. This issue is of critical importance for the feasibility of the evaluation, for ownership of the process and its results, for transparency, and for potential interest in and use of the findings. Donors generally carry out their conflict prevention, peacebuilding and statebuilding actions in support of and in partnership with host governments. A logical extension of that co-operation is working together in evaluation. Such partnerships, however, may pose challenges where governments lack legitimacy or are primary actors in an ongoing conflict. The political context and its high stakes not only affect external partners, they are also likely to have very real impacts on

how and why partners engage in an evaluation process. Donors need to learn more about managing partnerships in evaluations in fragile, conflict-affected settings.

Considering a joint evaluation

Joint evaluations that bring together different actors (development agencies, partner countries, etc.) can contribute to harmonised approaches where analysis and follow-up are shared. They are also seen as promoting opportunities to generate additional learning about how a variety of activities "add up" to produce an overall development impact. Joint evaluations often have a broader scope, capturing a more complete picture of development co-operation in a particular context. Some will cover many – even all – interventions in a particular conflict and fragility zone to assess their combined impacts. For example, by pooling information from several partners, the multi-donor evaluation in Southern Sudan was able to cover some 85% of the entire donor portfolio (Box 3.6). The OECD DAC Evaluation Network's *Guidance for Managing Joint Evaluations* (OECD, 2006) contains a number of practical suggestions and details as to why joint evaluations are conducted in development and how to conduct them.

Box 3.6. **A joint evaluation in Southern Sudan**

A major joint evaluation of peacebuilding efforts in Southern Sudan brought together a large number of development partners along with representatives of what would soon be the independent Government of South Sudan. A reference group in Southern Sudan was established and chaired by the transitional Ministry of Finance and Economic Planning to oversee and interact with the evaluation team during the evaluation. This group also involved representatives of government institutions, donors and agencies, the United Nations, the Joint Donor Team, and the NGO Forum Secretariat.

The evaluation concluded that support to conflict prevention and peacebuilding had been only partially successful. Donor strategies did not fully incorporate key drivers of violence, which resulted in an overemphasis on basic services and a relative neglect of security, policing and the rule of law, all essential to state formation. Assistance in preparing Southern Sudan for secession was insufficient. There was an over-use of nominally "good" practice – particularly with respect to ownership and harmonisation – at the expense of much needed in-depth knowledge and field presence. While harmonisation, co-ordination and alignment do not run counter to conflict prevention and peacebuilding *per se*, they are not in themselves sufficient responses to state fragility.

Source: Bennett *et al.* (2010).

Writing terms of reference

The preceding steps should inform the development of an evaluation's terms of reference, which outline what is expected of the evaluator or evaluation team and from the evaluation process itself. As such, it is a key instrument for managing expectations of evaluations and helps guide the evaluation process. The terms of reference or scope of work is usually written by the person(s) commissioning or managing the evaluation. Terms of reference should specify whether or not a final report will be published and other requirements for completion of the process. It may be useful to describe in the terms of reference what dissemination is planned and who will be responsible for it. Annex C contains a sample terms of reference document.

Setting up evaluation management

Commissioners and planners should set out the procedures for managing an evaluation. They should state clearly who is responsible for what (headquarters, field office, evaluators, partners, etc.), the team's degree of independence, the role of the evaluation team leader, decision-making processes and procedure for the clearance of reports. Experience has demonstrated the importance of being clear about whom the evaluation team should report to and how, including criteria for acceptance of different reports (length, content, writing style, etc.). Additional critical questions to be addressed are:

- How will management of the process be handled and what will the role of the different parties be? How will relations with the partner government(s) and conflict parties be managed?
- Should there be a reference or steering group? A reference group is an advisory committee that serves as an intermediary between management and the evaluators. It may also provide independent oversight of the evaluation. The group is usually made up of a variety of stakeholders and experts and can be used as a way of including stakeholders who are not otherwise directly involved in the evaluation.
- Will there be a management group? Who will be involved? A management group is useful when conducting a joint evaluation – it is made up of a group of selected representatives from those donors and agencies commissioning the evaluation and is responsible for the management of the evaluation process, including contracting and overseeing consultants (on behalf of the other donors and agencies).

Calendars, deadlines, and funding should be clearly and realistically determined and sufficiently flexible to adapt to a rapidly changing context. To the extent possible, the evaluation management should ensure a balance of genders in steering, management and reference groups.

Selecting the evaluation team

An evaluation team made up of members with complementary skills tailored to the task ahead is recommended. Planners and commissioners should specify the required competencies in the terms of reference and, if relevant, the tender document. They should spend time on this task, as it will be decisive in the evaluation process. People who are knowledgeable about conflict and peace – and the different priority areas within the field relevant to the evaluation subject – are critical to the quality of evaluations of donor engagement in situations of conflict and fragility. However, it is equally important to have knowledgeable evaluation experts in the team. As stated above, performing a gender-sensitive evaluation may well require gender experts in the evaluation team.

Particular attention also needs to be given to the perception of bias in the team. When hiring staff from the conflict-affected region or conflicting groups, it is important to take into consideration and adjust to possible threats to them, the rest of the evaluation team, and the credibility of the evaluation itself, which their involvement may jeopardise. The risk of the team being perceived as biased, or not being given access to certain information, should also be weighed, and the report should address and describe any implications for data collection or analysis.

Controlling quality

Ensuring the quality of the evaluation process and products – including reports – is key. There are several ways to organise quality assurance. The *DAC Quality Standards for*

Development Evaluation (OECD, 2010) can provide guidance as to what to emphasise. In some cases, often when dealing with challenging or complex multi-donor evaluations, internal or external experts review draft evaluation products for quality. They can do so prior to sharing the products with the wider stakeholder group, which can, in addition to improving quality, save time and money by reducing the number of rounds of revisions with the wider group. Product quality assurance and approval may also be tied to payment schedules.

© OECD 2012

Chapter 4

Conducting an evaluation in situations of conflict and fragility

Chapter 4 considers the business of conducting an evaluation. It begins with the inception phase, and then looks at how to identify theories of change and the implementation logic underpinning the activity being evaluated. The next step is the issue of gaps in baseline data and how and where to source data in order to plug the gaps. The chapter looks at the criteria evaluators should use, focusing in particular on the DAC criteria of relevance, effectiveness, impact, sustainability and efficiency. It then describes how to bring an evaluation to a close. The chapter looks at drawing conclusions and issuing recommendations and at the reports evaluators produce. The next step the chapter discusses is communicating the evaluation's (positive or negative) results to stakeholders and disseminating the lessons learned. Finally, to close the loop, the chapter emphasises the importance of feeding findings back into programme design and management.

A number of core steps should be part of any evaluation. Evaluators and managers (and others involved in the actual process of evaluating) do not have to perform them in the order in which they are set out below. But they should bear them in mind.

Key steps in the evaluation process

- ✓ Allow for an inception phase
- ✓ Identify the implementation logic and theory of change
- ✓ Deal with missing baselines and other gaps
- ✓ Gather data
- ✓ Examine the effort using various criteria
- ✓ Draw conclusions and make recommendations
- ✓ Conduct reporting
- ✓ Ensure quality
- ✓ Feedback on the evaluation
- ✓ Management response
- ✓ Disseminate findings
- ✓ Feedback into programming and learning

Allow an inception phase

Given the complexity of conducting evaluations of donor engagement in fragile and conflict-affected situations and the underlying weaknesses in programming, an inception phase can help identify issues that need to be addressed before proceeding with the evaluation. The scope of an inception phase can range from simple to complex. Usually it will involve desk study, document review, and the production of an inception report, but some preliminary field work may also be required. An inception phase may be useful for assessing or conducting conflict analyses. It could also be used to map the evaluation subject, conduct a donor policy and portfolio or country policy analysis, or carry out thematic studies as needed.

Evaluators present the results of an inception phase in an inception report that can be the basis for discussing data availability, evaluability, and the feasibility of planned data collection strategies. Inception reports are also useful tools for adapting methods and for fine-tuning the approach chosen to address the key issues specified in the terms of reference. (See "Reporting" in this chapter).

Identify and assess the theory of change and implementation logic

A theory of change is a set of beliefs about how change happens. It explains why and how people think certain actions will produce the changes they desire in a given context, at a particular moment in time (Weiss, 1995; Church and Rogers, 2006). It is a term used to

describe the links between the context, the intervention inputs, the implementation strategy, and the intended outputs and outcomes. Implementation logic, also called programme logic, is a term that describes why an activity is doing what it is doing, usually at the project level. It is closely linked to the theory of change. Evaluation can reveal whether success, or lack of it, is attributable to programme design and theory or to implementation. Contributing to testing theories of change and implementation logic is one of the prime contributions evaluation can make to research and broader learning.

In situations of conflict and fragility it can be especially important for evaluators to identify theories of change, since they are often implicit, unexamined, and untested. In Kosovo, for example, the international community operated for several years on the assumption (theory of change) that peace could be achieved by improving relations between the two main conflicting parties. On the basis of this theory, it funded numerous programmes to promote dialogue, exchanges, youth interaction, women's groups, and so forth. All were aimed at building cross-community relationships. However, a study found that without intra-community work to bridge internal divides and create more responsible leadership, cross-community interventions had little effect. The activities were operating on an incomplete theory of change. The study also found that cross-community relationships did not help prevent violence or strengthen collective resistance to violence in 2004. Bonding social capital – intra-ethnic networks of trust and reciprocity – and mono-community leadership, motivated by the strategic desire to secure independence for Kosovo, led to more effective action (Chigas *et al*., 2006).

Causation: Between theory and outcomes

A theory of change can take different forms at different levels. Theories of change at the country and conflict levels may be quite general – for example, strengthening the capacity of the government will help improve governance and therefore reduce conflict and violence. In this broad vision, development actors may operate in different ways (e.g. through budget support or capacity support to individual ministries) or work in different sectors, such as social services or finance and planning.

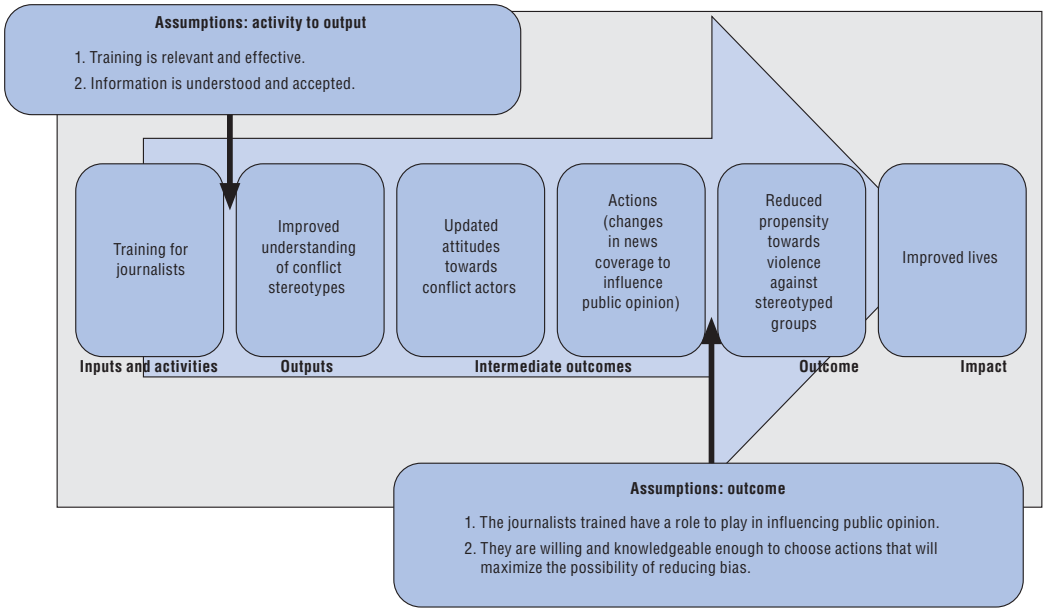
Evaluations of specific projects or programme would be likely to formulate more directly causal theories of change which link specific inputs and activities to desired micro-level outcomes and to broader peace dynamics. For example, in the Southern Sudan evaluation (Box 3.6), the team identified broad theories of change across the full portfolio of interventions they were examining. They did not drill down to the specific implementation logic level or look at individual projects. Instead, they grouped the multiple activities of different donors under broad theories of change, and then tested each one.

Programme documents do not always state in explicit cause-effect terms how a programme aims to produce its intended outputs, outcomes, and impacts. Accurate, clearly worded theories of change are necessary for effective programming and should be evaluated. More importantly, the assumptions that underpin theories should be the subject of evaluation. By identifying how an intervention was expected to contribute to sustainable peace or address conflict and fragility factors, evaluators use the theory of change to assess relevance, effectiveness, and impact. An example of this thinking is the imaginary anti-bias peace journalism programme described in Annex C and illustrated in Figure 4.1.

The programme could be understood to work on a theory of change whereby it will train journalists and increase their knowledge of conflict dynamics, so reducing bias in

reporting and contributing to a reduction in tensions and violence between warring groups. An evaluation of this programme would test the underlying theory that improved knowledge reduces reporting bias and the secondary theory that improved reporting lessens the tendency to resort to violence in the broader community. Key questions for analysis would be how workshops, awareness raising, and skills development actually change conflict reporting and what impact they have on critical conflict dynamics. The programme could track the language used in reporting before and after training. It might also survey public attitudes and, at the same time, the programme activities to see if they were achieving the expected results or whether unexpected obstacles had arisen. It might turn out, for instance, that individual journalists have very little influence over the use of inflammatory language and that editors determine the use of "colourful" language to boost sales, so reinforcing stereotypes. Such a finding would suggest that the "theory" of training journalists to influence public opinion was flawed.

Figure 4.1. **Assumptions underlying the theory of change in a fictitious peace journalism programme**



Source: Content from OECD (2008b); diagram adapted from Gaarder, 3iE (2011).

Where do theories of change originate?

One task related to identifying and evaluating theories of change is to identify where they originate. Are they empirical – based on solid evidence from past programmes? Are they drawn from programme designers' personal and professional experience or from the experience of the stakeholders and beneficiaries consulted during programme design? Or are they research-based?

When working theories are not explicit they will need to be uncovered through document review or interviews with implementing staff and other stakeholders. The evaluation process may also reveal that different stakeholders involved in a particular intervention operate on different assumptions (theories) about how their efforts will promote change for peace. By drawing out underlying theories and opening up a discussion of the validity of different theories, the evaluation process can be useful in helping those

involved in a programme reach consensus on what they are doing and why. This clarity can help improve programme design and implementation.

Evaluations of donor engagement in fragile and conflict-affected situations very often cover multiple projects and – in some cases – multiple donors. Such donor engagement may involve multiple theories or multiple variations of a broad general notion, e.g. more aid for social services supports stability.

Box 4.1. **Evaluating success and failure in peacebuilding**

When evaluating it is important to distinguish between and analyse various types of success and failure as an input to learning and future programme design. In this field it helps to distinguish a failure of the theory (wrong assumption about how change will happen) from a failure in implementation. Theory failure indicates the failure of a conflict prevention or peacebuilding activity due to a flawed causal relationship – in other words, underlying assumptions about how to bring about change in this context are false. A faulty theory of change could be based on an inaccurate conflict analysis, or it could reflect misdirected priorities or mismatched objectives.

Implementation failure denotes a problem with the execution of the activity itself (inputs/outputs, staff capability, timing, location, security environment or budget) or with management systems. Such problems could include sudden changes in the conflict that disrupt or reverse progress, despite an otherwise well designed activity.

Source: OECD (1999).

Gather data

As mentioned in the introduction, evaluators and evaluation managers sometimes encounter weaknesses in policies, strategies, and interventions. There may be unclear or unstated objectives, a poorly articulated theory of change or programme logic (as described above), missing indicators, no monitoring data, or no baseline information. The security situation can have serious implications for data and may particularly affect the ability of evaluators to travel to certain regions or countries and to access affected communities, programme beneficiaries or other key informants.

Evaluation teams, working in co-operation with the commissioning organisations, must address the issue of weak or missing data. They should consider ways to (re)construct or compensate for missing baselines and other data during the evaluation process, bearing in mind that nothing can ever fully replace solid planning. Tips on how to compensate for or work around such gaps without endangering the quality of the evaluation can be found in a number of publications (World Bank, 2005b; Bamberger *et al.*, 2006; OECD, 1999). Box 4.2 considers an example of how a donor-supported project in Afghanistan is working to increase the availability and quality of data.

A key phase in the evaluation process is collating data to trace the "story" of a programme or policy and its effects, which include how they affect stakeholders' viewpoints. Evaluators gather information from programme documents and reports, any available monitoring data, and interviews with programme staff, partner organisations, local officials, target groups, participants, third parties (groups such as neighbouring communities not targeted by a programme), and international and national actors who

Box 4.2. **Building statistics for monitoring and evaluation in Helmand Province**

The Helmand Monitoring & Evaluation Programme (HMEP) is a programme funded jointly by the UK's Department for International Development (DFID) and the Provincial Reconstruction Team (PRT). It is designed to assist improved delivery and effectiveness of stabilisation and development programming in Helmand, Afghanistan. HMEP collates data from third-party sources as well as drawing on a dedicated research capacity to establish baselines and monitor indicators of progress against the Helmand Plan. Data is presented graphically and geospatially and stored on an interactive database, accessible from an online website. HMEP produces quarterly monitoring and analytical reports and up to four *ad hoc* reactive reports per year aligned with PRT and DFID reporting requirements. In addition, HMEP is set up to support the development of programmatic capacity in the PRT and the integration of monitoring and evaluation into planning and programme implementation.

Source: Coffey International Development, "Helmand Monitoring and Evaluation Programme", *http://uk.coffey.com/our-projects/helmand-monitoring-and-evaluation-programme.*

observe the intervention from a distance (civil society organisations, research institutions, donors, media, academia, think tanks).

Data collection strategies depend on an evaluation's design, its purpose, and the information sources available. Strategies that can be used include random and purposeful sampling. Quantitative and qualitative data can be collected through censuses, observation, household surveys, interviews, questionnaires, anthropological or ethnographic research, participatory workshops and discussion groups. National statistics systems and major NGOs will often have available demographic data, though these may not be suited to the required sample size or detail. Moreover, some private polling companies have begun to specialise in household surveys and data collection to professional standards in conflict-affected areas. Box 4.3 gives examples of data sources and the evaluation literature provides further detail on data collection techniques.

Prolonged violence and situations of high tension pose significant data availability problems that often restrict evaluators' work. However, there might be more data available than is first evident in the planning phase and in programme documents. High staff turnover often shortens institutional memory. Moreover, different external and national actors are frequently poorly co-ordinated and do not always share the information they collect with other actors to whom existing studies, assessments, data, evaluations and surveys may not be known or available. Valuable baseline data might be gathering dust in documentary archives somewhere with few aware of it. It is sometimes worthwhile investing time up front in an evaluation phase to map data already collected, either as part of the evaluation inception phase or as a pre-study.

Evaluators and other researchers should share and use data collected by others, as too many are gathered and allowed to lie fallow. It is also advisable that those commissioning an evaluation and the evaluation team agree to contingency plans for sudden shifts in the situation on the ground that may impact collection of data that are necessary for the evaluation. Evaluation managers may include data management provisions in the terms of reference.

Box 4.3. **Examples of data sources from Somalia and Afghanistan**

Somalia: Using multiple data sources and accounting for data shortcomings

A 2011 evaluation conducted by DARA (an independent organisation specialised in humanitarian evaluation) looked at the effectiveness of the humanitarian response in Somalia. It drew on the following data sources:

- Literature on past assistance to Somalia and the contextual variables, and 300 web pages and relevant publications, and assorted documents.
- Semi-structured group and individual interviews with 489 stakeholders. Of these, 189 (112 in Kenya and 77 in Somalia) were carried out with a range of individuals involved in the response, who included representatives from UN agencies, the Red Cross and Red Crescent, international and national NGOs, local government actors, and donors. Women accounted for 24% of interviewees. There were group interviews with over 300 people from the affected population.
- Field observations in Kenya and Somalia, in camps for the internally displaced and Somali refugees.
- Online survey of former staff and key informants involved in the humanitarian response (response was low with only ten respondents answering).

In total, the evaluation team gathered 3 117 items of information but encountered the following challenges: high staff turnover, limited time for field work, insecurity, poor consistency of data, disaggregation gaps (limited availability of data disaggregated by age, gender, vulnerabilities and outcomes in the period under review). Safety concerns prevented international evaluators from visiting areas proposed in the terms of reference, although national evaluators were allowed to visit some.

North East Afghanistan: Creating new data

The more fragile a state is, the less data tend to be available. Evaluation teams often have to collect their own data. In the context of an evaluation of the impact of aid in North East Afghanistan, the evaluation team gathered data from various sources: The core of the data was collected with two surveys among 2 000 head of household respondents in 80 villages. Evaluators also collected data from other sources to create a profile for each community, containing information on the history, demography, ethnic composition, political and social organisation and resource endowment. In other settings, much of these data would be readily available from existing sources, such as a national census. In the Afghan context, however, evaluators had to collect these data themselves – which was possible only in the context of a multi-year research project. In order to collect information on major events and changes affecting the communities, a monitoring and reporting system was set-up covering 40 villages. Local correspondents completed semi-structured reports four times a year. The evaluation team collected additional qualitative data during field research visits. Finally, evaluators obtained data on aid flows from various international development organisations in the region. The data were then compiled and used for quantitative and qualitative analysis of the impact of aid on stability.

Sources: Polastro *et al.* (2011) and Böhnke *et al.* (2010).

During conflict and in its immediate aftermath, when mistrust is rife and most intervention stakeholders also have a stake in the conflict, the reliability (not to mention availability) of data and information provided is often particularly problematic. Depending on their own position within the conflict, different actors may have diverse, or even

contradictory, interpretations of an intervention's positive and negative impacts or relevance. To ensure reliability, evaluators should use multiple sources or types of information and a mix of sound quantitative **and** qualitative data. They should triangulate the data they use, ensure that sources are transparent, and verify the data's validity before analysing them – by fact checking with key stakeholders and interviewees, for example. By combining multiple data sources, evaluators can offset the bias that comes from relying on a single type of information and single observers. In the evaluation report's description of methods, any issues around data (including data gaps or problems with inconsistency) should be described along with the impact these data problems had on the reliability and validity of the evaluation's conclusions.

To avoid increasing tension between groups (in accordance with the "do no harm" principle), decisions about how to involve various groups in data collection should be based on a clear understanding of stakeholder roles and interests, drawing on the understanding of stakeholders developed in the conflict analysis. Where access or security concerns impinge on data gathering, other methods – including consulting with knowledgeable proxies able to provide representative views or information – could be explored. Consideration to coverage may be required – to include perspectives from both within and outside the capital city, for example – depending on what issues the evaluation intends to address.

Evaluations should draw on accurate, relevant data about women, men and gender relations. This helps ensure evaluations are gender and conflict sensitive, makes gender disparities more visible, and assists in answering the evaluation questions. Gender-disaggregated statistical data and gender analysis (information on the position of women and men in a particular conflict context) should inform the evaluation. Where disaggregated data are not available, combinations of quantitative and qualitative approaches can be used to gain the necessary understanding of gender relations and differentiate results for men, women, boys and girls. Where it is deemed that inclusion of such data is not necessary or not feasible, this should be clearly justified in the evaluation report. For example, in the case of an evaluation of the impact of aid on local perceptions in rural North East Afghanistan (Böhnke *et al*., 2010) the evaluation team used a survey. Though the team wanted to include women and youth, in the end the survey respondents were all male heads of households. Their decision was based on an analysis of the local context and consultation with regional and gender experts, and on an earlier experiment using female researchers to survey Afghan women, which found that security threats and the social custom of having men accompany each female researcher made it nearly impossible for the teams to work effectively. The evaluators decided that the choice not to include women in the survey was justified in this case because the opinions of older men (as the political representative of their household) were most relevant to answering the key evaluation questions. They verified the decision with experienced local teams and tested it before doing the initial baseline. In the evaluation report the team explained this decision, stating that, "while we think that the tasks at hand warrant these choices and the resulting limitations, we are nevertheless conscious of the fact that our research design is not equipped to capture trends for all of Afghanistan, nor does it capture perceptions of women [...]" (Böhnke *et al*., 2010). Transparency over methods and data limitations is a requirement of high-quality evaluations.

Many interventions work to build peace and prevent conflict by creating change in people's attitudes, thought processes, and relationships. In such cases, it may be necessary

to collect attitudinal data, conduct interviews, workshops, or focus group discussions with stakeholders, or carry out surveys to collect quantitative data. Measuring intangible changes in areas such as perceptions through interviews requires the same triangulation vetting as other types of data.

Relevant interviewees (primary sources) may be missing, because they are either in prison, dead, or unavailable in other ways. The evaluation report must account for such data gaps and make clear how they have been compensated for. Investing time early in the preparation phase for researching the data situation should help prevent surprises like missing data at a later stage, which can derail or delay the evaluation process.

Criteria for evaluating interventions

When evaluating development co-operation programmes and projects it is useful to consider the criteria outlined in the *DAC Principles for Evaluation of Development Assistance* and additional criteria that may be valuable. The analysis of these criteria forms the main content of an evaluation report. This section suggests how each OECD DAC evaluation criterion might be adapted to the field of conflict and fragility. The specific questions outlined below are guiding examples, not a comprehensive or obligatory list. The evaluation criteria are interlinked, each criterion shedding light on the intervention being evaluated from a slightly different perspective in order to develop as comprehensive a picture as possible of the intervention. When read together, the criteria should assist the evaluation team in developing a clear understanding of the activity or policy being evaluated and its contribution to statebuilding and peacebuilding. This adaptation of the criteria is based on contributions from Paffenholz and Reychler (2007), the draft guidance (OECD, 2008b), OECD and CDA (2007), and Anderson and Olson (2003).

Relevance

The relevance criterion is used to assess the extent to which the objectives and activities of the intervention(s) respond to the needs of beneficiaries and the peacebuilding process – i.e. whether they address the key driving factors of conflict revealed through a conflict analysis. Relevance links the outcomes of the conflict analysis with the intervention's objectives, although the relevance of the intervention might change over time as circumstances change. Understanding relevance may also involve an assessment of the extent to which an intervention ties in with overall strategies and policy frameworks of the country or external partners. Different conflict groups or actors may have different perspectives on the relevance of an intervention and its results. Women and men may also perceive the relevance of the intervention differently.

Assessing an intervention in relation to the conflict is key to evaluating its relevance. If staff, managers, or others involved in design and implementation have already carried out some kind of conflict analysis, its accuracy and use should be assessed. Assessing whether or not the conflict analysis proves (or has proved) accurate will be an important aspect of the evaluation. Assessing the analysis will contribute to learning and to refining theories about why violence occurs and what the most important determinants of conflict dynamics and long term statebuilding are. If no process of systematic analysis has taken place previously, the evaluation team may either talk to staff and stakeholders to understand what underlying (unarticulated) conflict understanding is guiding their work, or facilitate a more formal exercise for conducting a conflict analysis. The analysis is

needed to identify the expected connections between the programme outcomes and peacebuilding goals and the relevance to key driving factors of conflict and fragility.

Questions on relevance might include the following:

- Is the intervention based on a valid analysis of the situation of conflict and fragility? Has the intervention been flexibly adapted to updated analyses over time?
- In the light of the conflict analysis, is the intervention working on the right issues in this context at this time? Does the intervention appear to address relevant key causes and drivers of conflict and fragility? Or does it address the behaviour of key driving constituencies of the conflict?
- What is the relevance of the intervention as perceived by the local population, beneficiaries and external observers? Are there any gender differences with regard to the perception of relevance?
- Are the stated goals and objectives relevant to issues that are central to the situation of conflict and fragility? Do activities and strategies fit objectives, i.e. is there internal coherence between what the programme is doing and what it is trying to achieve? Has the intervention responded flexibly to changing circumstances over time? Has the conflict analysis been revisited or updated to guide action in changing circumstance?

Effectiveness

Effectiveness is used to evaluate whether an intervention has met its intended objectives with respect to its immediate peacebuilding environment, or is likely to do so. The key to evaluating effectiveness – and thus the linkage between outputs, outcomes and impacts – is finding out to what degree the envisaged results have been achieved and noting changes that the intervention has initiated or to which it has contributed. Furthermore, as most of the activities undertaken in situations of conflict and fragility are not explicitly geared towards sustainable peace, it is important to draw a distinction between two kinds of results. One is "programme effectiveness", i.e. to what extent the programme achieved its stated objective. The other is – if the programme met its objectives or goal – the immediate or secondary outcomes as they relate to peacebuilding and conflict dynamics identified in the analysis.

A theory of change analysis can be used to assess the criteria of effectiveness. This involves assessing whether an intervention is based on a sound theory and logic and whether these are proving (or have proven) to be true. The understanding of effectiveness is also linked to the conflict analysis. A programme or policy may do good or do well and still not change the underlying dynamics or key driving factors of conflict and fragility identified by the conflict analysis. Also, external factors, unrelated to and beyond the control of the activity in question, may be more significant factors of peace and conflict, in which case effectiveness will be understood relative to these broader dynamics and trends.

Box 4.4 describes how Germany evaluated the effectiveness of its civil peace programme using a framework that combined theories of change analysis with empirical research.

Box 4.4. **Evaluating the effectiveness of the German Civil Peace Service**

Germany commissioned an independent evaluation of its Civil Peace Service programme, which promotes non-violent ways of dealing with conflict through a variety of education, training and service activities. The evaluators developed a comprehensive evaluation framework that enabled systematic comparison of data across eight country cases. To compensate for insufficient data, a combination of evaluation approaches and methodologies was developed. In addition to reconstructing baselines and describing the Civil Peace Service's underlying theories of change, the evaluation team assessed the plausibility of outcomes, i.e. the likelihood that the intervention had achieved, or would achieve, the hoped-for outcomes. This was done by checking Civil Peace Service interventions against a list of conditions for effectiveness developed during empirical research. The evaluation team also identified good practice examples from among the interventions evaluated and then identified which factors contributed to their effectiveness.

Source: Paffenholz (2011).

Examples of questions that can help determine effectiveness include:

- Has the intervention achieved its stated (or implicit) purpose, or can it reasonably be expected to do so on the basis of the outputs and outcomes?
- Is the theory of change based on valid/tested assumptions? Are there alternative theories of change?
- To what extent did donors identify and adequately manage context-specific risks?
- Is the intervention achieving progress within a reasonable time frame, or will it do so? Is it possible to accelerate the process? Should the effort be slowed down for any reason?
- What major factors contribute to the achievement or non-achievement of objectives?
- Has the intervention achieved different results for women and men and boys and girls?

Depending on the activity being evaluated and the scope of the evaluation, some specific examples of conflict-specific questions on effectiveness include: Does the intervention prompt people to increasingly resist violence and provocation? Do the stakeholders affected have a significant impact on the situation of conflict and fragility? Are the right/key people or many people being addressed? Were gender equality and relevant horizontal inequalities (ethnic, religious, geographical, etc.) that drive conflict taken into consideration and what are the results? Does the intervention result in an increase in people's safety and in their sense of security? Does it improve non-violent forms of conflict resolution or power management? Does it result in a real improvement in relations among groups in conflict, demonstrated in behaviour changes?

Impact

The criterion of impact refers to the wider effects produced by an intervention. Such effects may be positive or negative, and may be produced directly or indirectly, intentionally or unintentionally. In fragile and conflict-affected contexts the criterion of impact is used to identify and evaluate the effects of the intervention on the key driving factors and actors of the conflict, as well as on broader development and statebuilding processes, as relevant. Assessment should cover both the desired changes the intervention aimed to achieve and any unintended (or unexpected) positive or negative results.

Changes in behaviour and attitudes, of the kind many peacebuilding interventions seek, are often difficult to measure and can take a long time to achieve. With this in mind, it may be too soon to reasonably expect significant impacts on conflict drivers (for example, for activities aimed at reforming institutions). In this case the evaluation might focus on outcomes and short term impacts and test the theory and programme logic to predict whether the current strategies are reasonably likely to contribute to peace over the long run.

Methods for evaluating impact are covered elsewhere in evaluation literature, for example in the guidance commissioned by the Network of Networks for Impact Evaluation (Leeuw and Vaessen, 2009). The rigorous quantitative methods associated with impact evaluation and randomised control trials are considered not feasible in many situations of conflict and fragility, (although useful experiments are underway at the programme level). Still, it is particularly difficult to apply such methods to high-level questions of peace and conflict across various interventions at country level or to assessments of overall donor engagement in a conflict setting. Where causality cannot be reliably determined using rigorous methods, evaluators may present plausible explanations for their conclusions regarding impact, though limitations should be made explicit. Box 4.5 describes quantitative impact evaluation methods in more detail.

When violent conflict is extreme, analysis may have to shift towards monitoring or "real-time" evaluation functions that focus on output indicators. They are easier to track and less prone to attribution failures than impact indicators.

Examples of questions that can help determine impact may include:

- What are the primary and secondary, direct and indirect, positive and negative, intended and unintended, immediate and long-term, short-term and lasting effects of the activity or policy in question? Does it exert a significant effect on key factors for conflict or peace?
- Drawing on the conflict analysis, what key drivers of conflict and fragility were affected and how? What changes can be ascertained in attitudes, behaviours, relationships or practices (of how many people and/or classified according to selected criteria such as gender)? Are there any secondary negative effects?
- How has the situation changed over time and what, if any, has been the contribution of the intervention to those changes?
- What impacts have the interventions had on specific indicators of well-being, such as health status or poverty levels, addressed by the intervention? What are the impacts on long-term development trajectories?
- Has the intervention impacted policy? How do these policies relate to the conflict?

Sustainability

Sustainability is defined as the continuation of benefits on end of assistance. In an environment of conflict and fragility, sustainability includes the probability of continued long-term benefits and resilience to risk over time, as well as lasting benefits in the economy, institutions, human resource management, etc. As in other fields, sustainability also includes "ownership" of the peace and development processes. Experience and peace research demonstrate that peacebuilding processes are long term and call for long-term engagement that can weather setbacks (OECD and CDA, 2007). In conflict-affected regions,

Box 4.5. **Quantitative methods to evaluate impact in settings of conflict and fragility**

An impact evaluation measures the net impacts of an intervention by comparing its results with a counterfactual – a measure of what would happen in the absence of the intervention. Although experience with the impact evaluation approach in peacebuilding interventions is limited, it is being used increasingly in the development context and, indeed, in conflict and fragile settings.

There are two main methods of conducting a counterfactual comparison. The first is the experimental approach or randomised control trial (RCT) evaluation. RCT evaluations need to be designed and initiated before the intervention begins. The evaluators randomly assign potential beneficiaries to "treatment" and "control" groups. The comparison of impacts between the two groups reveals the net impact of the intervention. This approach can be used to pilot an intervention to inform programme design. The second main method comprises "quasi-experimental" approaches, whereby statistical techniques are used to construct a control group that can serve as a counterfactual. Such techniques include various forms of matching, instrumental variables designs, regression discontinuity designs, and panel methods. These methods estimate the impact of an intervention by ensuring, to the extent that the data allow, that beneficiaries and non-beneficiaries are similar in all ways except for their treatment status. One such technique is propensity score matching, which estimates the net impact of an intervention by matching beneficiaries with non-beneficiaries who are similar to them in all other ways.

There are obvious drawbacks to these methods for evaluating peacebuilding interventions. Limitations include the frequent absence of baseline data, the desire to "treat" everyone (the humanitarian imperative), the political dangers of random treatment, and the qualitative nature of many relevant variables which are consequently less compatible with statistical methodologies. Impact evaluations are, nevertheless, often possible at the project or programme level and can provide valuable information on what works.

Source: Marie Gaarder (Norad Evaluation Department) and Annette Brown (3iE).

such engagement requires addressing those who have an interest in sustaining the conflict (sometimes called "spoilers" or the "hard to reach").

Questions regarding sustainability might include:

- Which steps have been taken or are planned to create long-term processes, structures, norms and institutions for peace? To what extent has the building of ownership and participation included both men and women?
- Will new institutions designed to address conflict and fragility survive? Are they being used? By whom? Does the intervention contribute to the momentum for peace by encouraging participants and communities to develop their own initiatives?
- Has a meaningful "hand-over" or exit strategy been developed with local partners or actors to enable them to build or continue their own peacebuilding initiatives?

Depending on the activity being evaluated and the scope of the analysis, some specific examples of conflict-related questions on effectiveness include:

- Does the effort result in the creation or reform of political institutions or mechanisms that deal meaningfully with grievances or injustices?

- Have those who benefit from and have a vested interest in on-going violence or instability, or who resist movement towards peace ("spoilers"), been addressed adequately?
- Will improvements in inter-group relationships persist in the face of new challenges and risks?
- Will the parties to a negotiated agreement honour and implement it?

Efficiency

The efficiency criterion is used to assess how economically resources (funds, expertise, time, etc.) are converted to results. In a conflict context, costs associated with prevention work will often be compared with the estimated costs of war or an outbreak of violent hostilities. Averted conflicts, however, are invisible. So there are, unfortunately, not many counterfactuals that compare the efficiency of prevention and conflict – though historical data may be used to reasonably estimate costs.

In addition to comparing the costs of supporting peace with those of war, evaluation could look at priorities in relation to key conflict drivers. It could ask this question: "Is this particular way of working against violence the most efficient option?" In a setting of scarce development resources evaluations should shed light on whether a particular activity is "worth it" compared to other actions or no action. Box 4.6 looks at the security costs of operating in a conflict zone, namely Afghanistan.

Box 4.6. **Security costs**

A joint evaluation of humanitarian aid and reconstruction assistance delivered in Afghanistan between 2001 and 2005 states that there was an overall security overhead of approximately 20% – i.e. operating costs were 20% higher in Afghanistan than in more stable settings for similar types of projects. The evaluation pointed out that these unexpected (or underestimated) costs made the Afghanistan programmes considerably more expensive than similar ones elsewhere. Further evaluative analysis and other data would be needed, however, to determine whether or not they may still be considered efficient.

Source: Chr. Michelsen Institute (2005).

When looking at development or humanitarian interventions in a particular conflict-affected area, evaluators should determine efficiency in comparison to other options for supporting development and peacebuilding in this or similar contexts. Questions on efficiency might include:

- Does the intervention deliver its results in an efficient manner compared to the counterfactual?
- How well are resources (human, financial, organisational) used to achieve results?
- Are there better (more efficient) ways of achieving the objectives?
- What was done to ensure the cost efficiency of the intervention? Did the intervention substitute local initiatives or did it come in addition to local initiatives?

Other criteria: Coherence and co-ordination

Much of the peacebuilding, statebuilding and aid effectiveness literature – and, increasingly, government policies themselves – points to the need for more coherent, better co-ordinated approaches. Funding for a particular conflict prevention project or peacebuilding initiative can be overshadowed and contradicted or – conversely – supported and sustained by other interventions that the same government(s) may undertake. However, the evidence base regarding the use and value of such approaches remains weak. An evaluation's terms of reference may therefore ask evaluators to look at this broader policy context and ask to what extent the activities are coherent with other policies or actions by other parts of government. Evaluations may also look at the degree of co-ordination among or between donors, government and other actors.

Evaluations may assess the costs and benefits of investing in co-ordination and coherence and any unintended consequences arising therefrom. It is important that co-ordination with other actors is not automatically presumed to contribute to better results. Potential risks that have been associated with more co-ordinated donor approaches include: heavy management structures which reduce flexibility and increase overhead costs; strong pressure on national stakeholders as a result of external actors' adopting a block approach ("ganging up" effects); and inappropriate influences on the neutrality, impartiality, and independence of humanitarian and development actors through association with military actors (shrinking humanitarian space). The evaluators must carefully examine the reality of co-ordination mechanisms in relation to the conflict analysis and other criteria (particularly relevance, effectiveness, and efficiency. Bennett *et al.* (2010) discovered in Southern Sudan that, in certain cases, single donor projects were more effective than co-ordinated pooling mechanisms.

Examples of questions that can help assess coherence and co-ordination might include:

- Were coherence and co-ordination factored into inputs and outputs (in other words, were they budgeted for and are they explicitly listed as outputs)?
- What were the roles of and relationships to other actors?
- What aspects of policy were successfully co-ordinated or made more coherent?
  - Joint analyses or understandings of conflict and fragility in the context?
  - Common development of priorities for funding or intervention?
  - Elimination or reduction in duplication of programming?
  - Useful divisions of labour across sectors, issues or problems?
  - Joint evaluations of programming at a country or sectoral level?
  - Useful sharing of experience and generation of lessons?
- Did co-ordination and coherence result in improved effectiveness, efficiency or impacts?
- How much time and what resources were spent on co-ordination? Was it efficient (cost *vs.* benefits) and appropriate? Did it reduce transaction costs?
- How was co-ordination achieved? What were the main constraints and challenges? Is it replicable in other situations (what contextual factors influenced co-ordination and results of coordination efforts)

Draw conclusions and make recommendations

The ultimate goal of any evaluation is to answer the key questions asked and present the results in a useful way. The evaluation report should clearly and credibly demonstrate the link between findings, conclusions and recommendations. Analysis of the collected data constitutes the basis for conclusions. Readers should be able to follow a clear line of evidence that supports the conclusions. Any recommendations should be well founded, relevant, targeted to the intended users of the evaluations, and actionable (meaning those who receive the information are able to act upon it) and should help improve the evaluated activity and any future activities. Evidence-based findings may be a useful basis for discussion between diverse actors in the conflict prevention and peacebuilding fields.

Over time, such conclusions will contribute to better understanding of how peacebuilding and statebuilding processes work, which can help inform strategies for intervention in conflict. Specifically, evaluations can add knowledge by demonstrating the accuracy (or inaccuracy) and use of conflict analyses, or by showing whether or not theories of change work and why. Box 4.7 gives an example of how an evaluation contributed knowledge with findings that yield insight into theories about the links between aid and people's perceptions of aid.

Box 4.7. **Can aid win over local communities? Testing the theory in Afghanistan**

In an evaluation of the impact of development aid on stability in rural communities in North East Afghanistan, the evaluation team examined two theories held by many development actors working in conflict zones. First, the assumption that stability in conflict or post-conflict contexts depends on whether the population perceives the state and its international partners to be legitimate and useful and so co-operates with them, or not. Second, the assumption that aid can influence local people's perceptions of the government and international actors. The evaluation team collected data on the perceptions of the rural population, and matched it with data on aid delivery to the communities. It was found that aid has little impact on how the population perceives international actors. It does, however, lead to more positive perceptions of the state.

Source: Böhnke *et al.* (2010).

Depending on the type of evaluation, conclusions and recommendations may be developed in a participatory format. For example, evaluators could present initial findings to a group of key stakeholders with whom they would then work to draw useful conclusions. On the other hand, if the focus of the evaluation is on accountability, the evaluators are likely to take a more non-participatory approach. Several recommendations from an evaluation of peacebuilding interventions in Southern Sudan are shown in Box 4.8.

Evaluators may discover major differences of opinion not only as to what happened, but as to the value of outcomes and impacts – particularly because individual and group understandings are highly determined by conflict and their own roles in it. Evaluation is, ultimately, a values-based exercise. What is viewed as a successful intervention by some groups may be seen as useless or harmful by others.

Box 4.8. **Making recommendations for donors and the Government of Southern Sudan**

The evaluation team in charge of assessing donor support to conflict prevention and peacebuilding in Southern Sudan (soon to be the independent state of South Sudan) from 2005 to 2010 (Bennett *et al.*, 2010) drew on the evaluation findings to formulate specific, targeted recommendations to donors and the Government of Southern Sudan. Concluding that donors did not adjust what they were doing on the basis of a sophisticated and nuanced analysis of power relations, causes of vulnerability, drivers of conflict and resilience indicators (conflict analysis), the evaluation team recommended that donors and the wider aid community:

- Ensure that revised and new programmes are always preceded by a conflict analysis that links wider dynamics to those specific to the area of operation.
- Plan, monitor and evaluate interventions according to the critical factors identified.
- Rate interventions on responsiveness to conflict factors.

The team analysed which conflict drivers should be prioritised and how this might be done. They suggested the Government of Southern Sudan (and its supporting donors):

- Allocate major resources towards creating and maintaining livelihood programmes for young men who are currently too easily drawn into criminal activity.
- Enable traditional authorities (chiefs) to address root causes of conflict (including disputes over land or bride wealth) at their customary courts by providing capacity-building programmes for these courts.

Develop effective oversight mechanisms to monitor security agencies.

Source: Bennett *et al.* (2010).

Reporting

Reporting takes place throughout the evaluation process (which includes the planning and inception phases). There are often three main reporting steps, though specific reporting requirements may be set by the commissioning agent and should be outlined in the terms of reference.

Inception report

After conducting a conflict analysis and gathering initial information, the evaluators will draft an inception report describing how the team intends to conduct the evaluation and answer the questions set out in the terms of reference. It presents risks and challenges, the methods to be used, data collection tools, indicators if relevant, operationalisation of the main questions, theories of change (implicit or explicit), sample selection tools, case studies (if not selected prior to commissioning the team), the structure of the report, and a work plan for the remaining work. Stakeholders usually comment on the inception reports, often as part of a reference group. (See Chapter 4 for a more detailed discussion of the inception phase.)

Draft report

A draft of the evaluation report is often circulated widely for comments and is a chance for stakeholders to comment on the evaluation. Sufficient time for comments should be built into the overall time frame. Some agencies may also require an "out-brief"

before the team departs from a field visit (between the inception and draft reports) to promote accountability, or other types of reports during field visits.

Final report

Though an evaluation may result in many different outputs, a written report is almost always completed. Reports and presentations will need to be translated into locally relevant language(s) to facilitate sharing with all stakeholders. The final report is sent to stakeholders. Target groups for dissemination should be agreed on at the beginning of the process. Many organisations now use the Internet as an alternative means of publishing the final report, either in part or in its entirety. In all cases, the confidentiality and safety of those who contributed to the evaluation should be carefully protected.

Management response and follow-up action

The recommendations and conclusions of the evaluation should be systematically responded to. The people or institutions targeted (generally donor agencies, implementing agencies or national governments) by each recommendation will respond and take relevant action. The terms of reference should include the process for reporting and responding to the evaluation. A formal response and follow-up by those in charge of a programme will often be required. To write the management response, to disseminate findings and lessons, and to engage in a learning process is the responsibility of the commissioning agency. The form of the management response and follow-up action required varies by institution and may be different for different types of evaluations. A response will often include an assessment of the quality and limitations of the evaluation.

In the case of a joint evaluation, a joint response or multiple individual responses should be used, depending on institutional requirements. In Box 4.8 the Canadian International Development Agency (CIDA) published a formal response to the findings of an evaluation of interventions in what was then Southern Sudan.

Box 4.9. **Canada responds to the evaluation of peacebuilding support in Southern Sudan**

The Multi-donor Evaluation of Support to Conflict Prevention and Peacebuilding Activities in Southern Sudan (Bennett *et al.*, 2010) made recommendations to the Sudanese Government and its international partners, including Canada. Canada has a "whole of government" approach in South Sudan, so relevant departments discussed the report findings. In addition, the Canadian International Development Agency (CIDA) Sudan Programme responded to the evaluation by publishing a formal management response. The response states that CIDA found the evaluation report to be a useful analytical piece. It also pointed out that the evaluation provided a unique learning opportunity because it was the first programme-level multi-donor evaluation in which CIDA had participated.

Box 4.9. **Canada responds to the evaluation of peacebuilding support in Southern Sudan** *(cont.)*

The managers described what they had done or planned to do regarding the specific recommendations. For example, on the recommendation to use better conflict analysis, the managers stated that they "[recognised] the importance of conflict analysis throughout the project lifecycle" and described how they undertook initiatives to promote conflict-sensitive programming, including conflict mapping to guide the design and implementation of new projects relating to children and youth and food security. Specifically, the design of projects together with the United Nations Children's Fund and the Food and Agriculture Organization with the United Nations Industrial Development Organization involved a two-day workshop on a collective mapping exercise to identify potential geographic locations for planned projects and to determine how to address root causes of conflict. Factors related to early recovery needs and opportunities for synergy between food security and youth development programming were taken into consideration.

Source: Canadian International Development Agency (2011).

Disseminate findings

Plans for follow-up and dissemination of lessons learned should be implemented, as agreed during the evaluation planning process described in Chapter 3. Appropriate means of communicating the results to specific target group(s), such as managers, staff and decision makers, should be used. Evaluation managers are increasingly using short summaries, local language translations, policy briefs, video clips and other tailored communication tools to reach different audiences with relevant findings.

Sharing the outcomes of an evaluation can be difficult when the results are perceived as negative or when they question strategies or approaches to which practitioners feel strongly committed. Stakeholders may resist questioning the effectiveness of their approach. Receptivity can be enhanced by emphasising the learning aspects of evaluation – and engaging stakeholders early on.

Feed back into programming and engage in learning

A completed evaluation should feed back into the early stages of planning and programme design and help to address challenges by providing more evidence on the validity (or not) of theories of change and data for comparison and reference. Evaluations carried out while a policy or programme is still going on can be used to adjust or redesign it.

In addition to the immediate use of the evaluation findings (by those commissioning the evaluation, for example), opportunities may be identified to feed evidence, lessons or broader conclusions of the evaluation into other policy forums and research activities. For instance, an evaluation of a particular programme working with women peace mediators may yield insights into perceptions of women in different conflict-affected communities which could be valuable to working in the conflict context concerned. Alternatively, it may have lessons for the effective implementation of mediation programmes and contextual factors for success which could be relevant to people working on mediation in quite different contexts.

Evaluators must, in particular, be transparent about the external validity of evaluation findings – i.e. how applicable they are in other contexts. The consequences of applying findings in the wrong way or in a context where they are not relevant can be very negative. This is one reason why it is important for evaluators to explicitly describe weaknesses and strengths of data and methods used, and how they impact findings, conclusions and recommendations.

Having completed the evaluation and learning process, decision makers, managers and staff should be better able to understand and improve strategies, outcomes and impacts, so making more lasting contributions to peace.

© OECD 2012

ANNEX A

Conflict analysis and its use in evaluation

Introduction

This guidance suggests the use of conflict analysis in planning, managing, and evaluating conflict prevention and peacebuilding programmes and policies. Conflict analysis helps to identify what is needed to address the conflict and to understand the context in which an intervention is to be implemented. As such, many practitioners will already be familiar with the use of this tool in designing projects and programmes. This annex seeks to further explain the role of conflict analysis in the context of evaluation.

A variety of conflict analysis approaches or frameworks are available and these are often used in combination with each other. The choice of approach will depend on the purpose of evaluation and the actors involved (some development agencies have institutionalised a particular type of conflict analysis, for example). While different in approach and coverage, most frameworks take the user through similar steps to identify the causes and drivers of conflict and fragility: examine the key stakeholders (actors and groups) who are affected by or influence how a conflict develops; understand the multifaceted context in which conflict and peacebuilding takes place – including state society relations and political economy; and assess the dynamics of a conflict, how it might evolve in the future, and what opportunities there are for preventing escalation (International Alert, 2007a).

Conducting or reviewing a conflict analysis for an evaluation

Evaluation teams are primarily concerned with conflict analysis from three perspectives. First, in assessing relevance, which includes the use of conflict analysis by managers or policy makers in determination of priorities or programme approach. Second, in order to assess the impacts of policies or programmes, the evaluation team needs to understand the conflict that those programmes and policies are attempting to influence or change. An evaluation team thus needs to understand the different approaches to, and tools for, conflict analysis in order to review the analysis performed at the design stage or conduct its own analysis. Finally, evaluators use the analysis to ensure their process is conflict sensitive.

Choosing the appropriate kind of conflict analysis

If the intervention being evaluated did not use a conflict analysis in the design phase, or if the analysis is implicit, or if it is not clear how the conflict has evolved since the outset/implementation of the programme, the evaluation team will need to obtain or undertake one itself to serve as the basis of the evaluative assessment. The level of effort

and resources required for this work should be included in the terms of reference and adapted to the scope of the evaluation questions.

There are many different models and frameworks for conflict analysis used by development donors and others engaged in working in and on conflict and fragility. Some models are formal, others informal. Informal analysis generally dominates where political sensitivity is high. "Traditional" models of conflict analysis focus on understanding the context and causes and on understanding the conflict's stakeholders and actors and their interests. Some widely used conflict-analysis methods do not help identify priorities or factors that are important to the conflict and fragility dynamic (OECD and CDA, 2007). Furthermore, some conflict-analysis tools produce a static snapshot – often in the form of lists of factors – without much sense of how they work together. The frameworks and tools that treat conflicts as complex systems and those that explore future scenarios point to ways of addressing this problem.

Evaluation managers and evaluation teams might consider a few questions in deciding what tool or combination of tools to use:

Purpose

- Does the tool provide sufficient information on causes, actors, dynamics and the context to assess the relevance of the activity to the needs of the peacebuilding process?
- Does the tool provide information on the appropriate issue areas, at the appropriate level and depth, to help evaluate the effectiveness and impacts of the programme or policy?

Assumptions

- Do the evaluators share the underlying assumptions about the conflict that form the basis for analysis? Is the tool's understanding of or assumption about the nature of conflict appropriate to the specific context in which the programme or policy is being implemented?
- Does this perspective correspond to the mandate and values of the organisation being evaluated?

Methodology and resource implications

- Does the tool's proposed methodology match the purpose of the analysis?
- Does the tool's proposed methodology agree with the ways of working of the evaluation team?
- Does the evaluation team have the capacity (skills, expertise, access, etc.) to use the tool well?
- How long does it take to produce a reliable conflict analysis?
- What are the resource implications of the selected tool (staff time, travel, seminar costs, facilities, data management)?
- Is the evaluation team able to allocate or secure the required resources?

Table A.1 outlines a few conflict-analysis tools developed and used by governments, multilateral agencies, research institutions and NGOs. It is not an exhaustive list, but is intended to give a sense of the variety of tools and range of approaches. Further conflict-analysis resources are listed in the bibliography. There is also a useful overview of conflict analysis compiled online at *www.conflictsensitivity.org*.

Table A.1. **Summary of selected conflict analysis tools**

Purpose	Potential users	Assumptions	Methodology and effort	Evaluation application
Conflict Assessment Framework – USAID				
Country and programme strategic planning to identify and prioritise causes of conflict based on understanding of impact.	Donor desk officers, implementing partners, mission staff, embassy staff, other government officials.	Pulls together best research on causes, level and nature of conflict to identify windows of opportunity.	– Combination of desk study, in-country visits, workshops and interviews. – Includes significant staff time: about 2 months.	– Relevant to both conflict sensitivity, prevention and peacebuilding. – Quality may vary depending on robustness of methodology used to gather data.
Conflict-related Development Analysis – UNDP				
Conflict, related programme planning and review aimed to understand linkages between development and conflict, increasing positive impact of development efforts.	Development agency staff and donors working in situations prone to and affected by conflict.	– Conflict caused by combination of security, political, economic and social causes and actor interests. – Development can cause violence.	– Data collection and analysis followed by workshop or expert study to analyse current responses and suggest ways forward. – Effort depends on method for data collection.	– Development-focused and linked to programming. – Useful at country or sector-level, less at micro level. – Quality of analysis depends on rigor of data collection.
Manual for Conflict Analysis – SIDA				
Country/ programme/ project planning to improve effectiveness of development co-operation and humanitarian assistance in areas affected by violent conflict.	Development agency staff, implementing partners.	Conflicts driven by structural instability, struggle for power and influence, and mutual fear and insecurity.	– Desk study, consultations and workshop to consider programme implications – Local ownership of analysis important - 6-12 weeks, pending scope of desk study.	– Focus on different levels of programming. – Relevant both for conflict sensitivity and planning at country and sector levels – No methodology.
Aid for Peace – Paffenholz and Reychler (2007)				
Assess peace and conflict relevance, risks and effects of development and humanitarian projects or programmes.	Development and foreign ministry officials.	– Examines both conflict and peace factors. – Framework for analysis of peacebuilding deficiencies and needs, conflict risks and effects of intervention on conflict.	– Desk study/survey of other interventions; field mission with 3-5 day training and workshop. – Potentially time consuming and costly, pending time for baseline study and mapping and number of field visits and workshops.	– Addresses both conflict sensitivity and peace and conflict programming. – Provides specific guidance on integrating peace and conflict lens into evaluation.
Making Sense of Turbulent Contexts: Analysis tools for Humanitarian Actors – World Vision				
Aims to improve ability to analyse dynamics of conflicts to impact programme and project planning and advocacy in emergency situations.	NGO emergency response, development and advocacy staff.	– Focus on chronic political instability, not just violent conflict. – Sees conflict as cyclical with periods of peace followed by conflict.	– Collection of tools to analyse actors, symptoms and political economy of conflict, generate future scenarios, and analyse strategic and operational implications. – Effort pending on scope of data collection and workshop.	– Focuses on macro level; how conflict will affect programme in future. – Flexible and adaptable to specific contexts. – Can be used for analysis of clusters of countries.
Conflict Prognosis: A Conflict and Policy Assessment Framework – Goor and Verstegen (2000), Clingendael Institute				
Aims to link early warning to policy planning and implementation.	Donor and embassy staff involved with foreign policy and development issues.	– Focus on indicators of internal conflict and state failure. – Uses Fund for Peace's measures for sustainable security as goal.	– External research and analysis to track indicators and identify problem areas and responses for workshop discussion. – Effort depends on size of workshops, and consultant involvement.	– Not programme specific, but focuses on broad policy or programme development. – Facilitates clarity on developments and trends, not causes.

ANNEX B

Understanding and evaluating theories of change

What are theories of change?

Theory of change is a flexible approach meant to encourage critical thinking in the design, implementation and evaluation of development activities. As described by Vogel (2012) "theory of change thinking" is being increasingly used in international development by a wide range of actors. This guidance encourages questioning strategies and activities that impact on peacebuilding and conflict prevention. It offers theories of change as one way to help evaluators assess and programme managers and decision makers think through the hypotheses of change and assumptions that underpin their work. Vogel describes theory of change as a process of analysis and learning that produces insight to support critical thinking throughout the programme cycle. It is also a flexible approach that may be helpful in encouraging innovation in programme strategies to respond and adapt to change in the context.

Aid work in relation to conflict and peace is often based on approaches, strategies and tactics that are rooted in theories of change (understandings about why particular inputs or activities are expected to achieve intended results [outputs, outcomes and impacts]) that are unstated or ill-defined. They are embedded in the skills and approaches of individual practitioners and peacebuilding organisations, their capacities and technologies, attachments to favourite methodologies, and the perspectives of different stakeholders about conflict and peace.

In the imaginary example of an anti-bias peace programme for journalists in Annex C, one question would be how the planned workshops, consciousness raising, and skills development might actually change conflict reporting. The programme could track the language used in reporting before and after the effort and also survey public attitudes. At the same time, it could see whether the activities were achieving the expected results – or if unexpected obstacles appeared. For instance, it might turn out that individual journalists have very little influence over the use of inflammatory language and that editors determine the use of such language to boost sales. That outcome would suggest that the "theory", about inducing changes in reporting by training journalists, was flawed.

One related task is to identify the sources of theories. Are they a) based on experience (the programme designers' personal and professional experience or that of the stakeholders and beneficiaries consulted during programme design); or b) research-based? Evaluation can contribute to improving the design and implementation of ongoing programmes. It can also uncover whether success, or lack thereof, is due to programme design and theory or to programme implementation.

A useful first step in enhancing strategies in conflict prevention and peacebuilding programming and evaluation is to become more explicit about underlying assumptions of how change comes about – that is, theories of how to achieve peace. Conflict prevention and peacebuilding activities are carried out on the basis of specific ideas and goals as to what they hope to achieve. Programme decisions are based on a number of factors – including assumptions about how to bring about peace and theories about how to bring about change. Peace practitioners select methods, approaches and tactics that are rooted in a range of theories of how peace can be achieved in a specific context. It is important to uncover these theories of change, both in order to test the theories against the realities of the conflict and to provide the basis for evaluating progress towards related objectives.

Theories of change in peacebuilding include those presented in Table B.1, though a systematic inquiry into ongoing and past conflict prevention and peacebuilding work would likely reveal many other theories underlying peacebuilding programmes.

Some theories focus on **who** needs to change – that is, which individuals and groups in society or which relationships need to change. Other theories concentrate on **what** needs to change. It may be an institution, a policy, a social norm. Still other theories are tied directly to a particular methodology or approach: **how** the change could or should happen.

Evaluating conflict prevention and peacebuilding theories of change

The impacts, effectiveness, relevance, efficiency and sustainability of a conflict prevention and peacebuilding activity rest to a large extent on the accuracy of its underlying theory of change. A false or incomplete theory may be a key explanatory factor for the failure of a programme, project, or policy. In contrast, good theories (based on an up-to-date, thorough conflict analysis) contribute to effective conflict prevention and peacebuilding action and successful interventions. Analysis of the theory of change is therefore a key aspect of any conflict prevention and peacebuilding evaluation. The pertinent theory should be reviewed in the evaluation report and be covered in the evaluation's findings, conclusions, and lessons learned. Such analysis will help contribute to a more refined understanding of how to bring about change for peace.

When conducting an evaluation, the evaluator or evaluation team should ascertain the theories of change of the peacebuilding intervention in question. While they are often variations on the generic theories presented in Table B.1, they should – for the purpose of evaluation – be reframed in relation to the specific context and using the intervention's particular terms.

At times, the theories in operation are obvious, even if unstated, in programme proposals and other documents. More often, the theories need to be uncovered through interviews with implementing staff and other stakeholders – or can be confirmed by those discussions. The evaluation process may also reveal that different staff members are proceeding on different assumptions (theories) about how their efforts will promote change towards peace. The evaluation process itself can thus be useful for helping to clarify this important dimension of intervention strategy.

The two real life examples which follow illustrate these points.

Example 1: Evaluating grassroots conflict prevention in Liberia

In the wake of the 14-year civil war in Liberia, a large international NGO received donor funding to develop Community Peace Councils (CPCs), a community-based mechanism for

resolving a range of disputes, with an explicitly inter-ethnic approach. The CPCs were designed to promote greater democratic participation through leadership development. The evaluation team first identified underlying theories of change and programme assumptions (derived mainly from discussions with local and international staff members) for the CPCs:

Theory #1: Establishing a new community-level mechanism for handling a range of dispute types will contribute to keeping the peace and avoiding incidents that have the potential for escalating into serious violence.

Theory #2: Creating inclusive structures for community problem solving can improve communication, respect, and productive interactions among subgroups in the community, and improve the access of disenfranchised groups to decision making.

Theory #3: Creating a new leadership group infused with democratic concepts and provided with critical skills can foster more effective and responsive leadership.

The evaluation team then discussed whether and how these theories of change were appropriate for the situation in Liberia, and how they were playing out in the programme. To begin, the team conducted an updated conflict analysis, based on interviews and focus groups with a wide range of people in the communities themselves. It then examined whether the programme was having the effects envisioned in the theory of change. For example, the team examined what kinds of conflicts the CPCs handled, and whether those conflicts had the potential for escalating and inciting widespread violence. If they did, then the CPCs would directly contribute to stopping a key factor in violent conflict. If, however, those conflicts were unconnected to the driving factors of the conflict, or the local-level conflict-handling mechanisms were not able to address the types of conflict most likely to escalate, then the CPCs would make little or no contribution to "peace writ large".

The evaluation team found that the CPCs were, for the most part, **not** handling the most serious and volatile disputes, which concerned land issues. The team then explored whether this was due to a failure in programme implementation or, alternatively, a theory of change that was incomplete or inaccurate. The main conclusion was that, while the CPCs were well set up and trained well, the CPCs became mostly excluded from handling land issues as communities were repopulated and traditional leadership patterns re-established. At the same time, the hope (and theory) regarding alternative leadership models proved unfounded, as traditional leaders gained control over the CPCs or used them to address issues they preferred that someone else deal with. The evaluation recommended that the agency work to expand the mandate and capability of the CPCs for handling land disputes by connecting them to land commissions and other emerging government structures. It should also be said that the CPCs did represent a useful developmental advance, even if they were unable to fulfil, as completely as hoped, a contribution to "peace writ large".

Example 2: The impact of international peacebuilding policies and programming in Kosovo

CDA Collaborative Learning Projects performed an extensive study of the reasons for the recurrence of inter-ethnic violence in Kosovo in the spring of 2004 and the relationship of that violence to policies and programmes undertaken by the international community. Among other things, the study identified the theories of change underlying the various approaches to improving ethnic relations. As is often the case, these underlying theories

were strongly influenced by the policies and (unspoken) assumptions of the international community. The multiple aid and development programmes were directly linked to implementation of internationally-established "Standards for Kosovo" and widely held beliefs regarding refugee returns, inter-ethnic relations, and a future multi-ethnic state as the basis for peacebuilding.

The Kosovo example concerns many agencies and multiple programmes. The study identified major programming approaches, and associated theories of change, some of which are listed here, and then examined the effectiveness of each, and their relationship (if any) to preventing violence.

A. Inter-ethnic and inter-religious dialogue

In Kosovo, the bulk of what agencies and community members identified as peacebuilding was labelled “dialogue.” Dialogue encompassed a wide range of activities: from social contact to structured conversations about identity and the promotion of mutual understanding, to problem-solving related to concrete issues, and to negotiation and mediation of agreements on land use in the Municipal Working Groups on Return. The most frequent theories of change for dialogue efforts in Kosovo were:

Theory #1: Involving Kosovar Serbs and Albanians in mutual discussions can help develop the conditions for the safe, successful and peaceful return of internally displaced persons to their homes. This, in turn, will promote reintegration, stabilisation of the environment and will reverse one of the negative consequences of the conflict.

Theory #2: Engaging community members in participatory decision making and the implementation of development activities can strengthen community relationships.

Theory #3: Promoting co-operation across ethnic lines regarding non-political issues of common interest (HIV/AIDS, drug use, business and entrepreneurship, women’s rights, infrastructure, etc.) can build stronger inter-ethnic ties and understanding.

B. Training and peace education

Training in conflict resolution, human rights, nonviolent communication and related topics was done in many communities, and, with dialogue, was one of the most popular approaches to peacebuilding programming. Youth camps, peace camps, archaeological camps, art camps and many others were widespread, as were multi-ethnic programmes of technical training in computers, project management, marketing, and other technical or professional topics. To a lesser extent, school-based peace education programmes were developed and included human rights education and tolerance education for children.

Theory #1: Providing people with better skills for conflict resolution will increase the ability of communities to settle disputes non-violently and reduce the likelihood of violence.

Theory #2: If people talk and play together they will build relationships and break down stereotypes.

C. Multi-ethnic projects and institutions

Along with dialogue and training, joint (inter-ethnic) projects and institutions comprised a significant proportion of the peacebuilding programming in the communities that were included in the Kosovo study. Some of the projects were the outcome of or

follow-up to dialogue, aiming to take the communication and relationship building beyond mere talk.

Theory #1: Developing activities that provide economic benefits to both ethnic communities (economic interdependence) will give people incentives to resist efforts to incite violence.

Theory #2: Providing opportunities for people to work together on practical issues across ethnic lines will help break down mistrust and negative stereotypes, as well as develop habits of co-operation.

Theory #3: If people have jobs and economic stability, they will be less hostile to the other ethnic group.

D. Democratic governance and capacity-building

Many international donors, agencies and NGOs have implemented peacebuilding activities designed to strengthen municipal government institutions to support integration of minorities, better communication and dialogue, and sustainable returns. They work on the following theory:

If we can improve administration and service delivery and establish non-discriminatory policies, this will reduce inter-ethnic tensions and demonstrate the viability of a multi-ethnic Kosovo.

Many programmes and policies integrated **several approaches and theories of change**. For example, a programme to facilitate returns of Kosovo Serb minorities included activities and approaches that reflected a combination of different theories:

- dialogue between the host community and returnees was facilitated on the assumption that dialogue would allay fears and re-establish relationships that would allow returnees to return to their homes in peace (Theory A #1);
- multi-ethnic committees to decide community priorities for development aid (Theory A #2);
- provision of equipment and seeds to a multi-ethnic agricultural co-operative (Theories C #1, 2).

Once the theories had been identified, they could be assessed in relation to the driving factors of conflict and the factors contributing to the absence of violence in some places in March 2004. The Kosovo study identified patterns of inter-ethnic violence and factors that contributed to the prevention of inter-ethnic violence – through extensive interviews in communities, some of which experienced violence in March 2004 and some which did not. The team then examined the programming approaches and any relationship to the factors that helped communities avoid violence. The study found that the failure of peacebuilding programming to achieve desired impacts was due partly to faulty theories of change and partly to problems in programme design and implementation.

Design problems included failures in the participant selection processes, fragmentation of programming, insufficient follow-up and limited resources for "soft" aspects of programming. As for implementation strategy, returnees were not central actors with respect to violence, although they were important victims of the conflict. The channelling of aid to returnees and to communities that accepted them, it turned out, prompted resentment and led to increases in inter-ethnic divisions rather than improved relations between groups.

In part, the theory of change on which the programming was based was faulty. With respect to inter-ethnic dialogue between host communities and returnees, the study found

Table B.1. **Common theories of change**

Theory of change	Examples of methods
Individual change: If we transform the consciousness, attitudes, behaviours and skills of many individuals, we will create a critical mass of people who will advocate peace effectively.	Individual change through training, personal transformation or consciousness-raising workshops or processes; dialogues and encounter groups; trauma healing.
Healthy relationships and connections: Strong relationships are a necessary ingredient for peacebuilding. If we can break down isolation, polarisation, division, prejudice and stereotypes between/among groups, we will enable progress on key issues.	Processes of intergroup dialogue; networking; relationship-building processes; joint efforts and practical programmes on substantive problems.
Withdrawal of the resources for war: Wars require vast amounts of material (weapons, supplies, transport, etc.) and human capital. If we can interrupt the supply of people and goods to the war-making system, it will collapse and peace will become possible.	Campaigns aimed at cutting off funds and national budgets for war; conscientious objection and/or resistance to military service; international arms control; arms (and other) embargoes and boycotts.
Reduction of violence*:* If we reduce the levels of violence perpetrated by combatants and/or their representatives, we will increase the chances of bringing security and peace.	Ceasefires; creation of zones of peace; withdrawal or retreat from direct engagement; introduction of peacekeeping forces and interposition; observation missions; accompaniment efforts; promotion of non-violent methods for achieving political, social and economic ends; reform of security sector institutions (military, police, justice system/ courts, prisons).
Social justice: If we address the underlying issues of injustice, oppression/exploitation, threats to identity and security, and peoples' sense of injury/victimisation, it will reduce the drivers of conflict and open up space for peace.	Long-term campaigns for social and structural change; truth and reconciliation processes; changes in social institutions, laws, regulations, and economic systems.
Good governance: Peace is secured by establishing stable and reliable social institutions that guarantee democracy, equity, justice, and the fair allocation of resources.	New constitutional and governance arrangements and entities; power-sharing structures; development of human rights, rule of law, anti-corruption; establishment of democratic, equitable economic structures; economic development; democratisation; elections and election monitoring; increased participation and access to decision making.
Political elites: If we change the political calculus and perception of interests of key political (and other) leaders, they will take the necessary steps to bring peace.	Raise the costs and reduce the benefits for political elites of continuing war and increase the incentives for peace; engage active and influential constituencies in favour of peace; withdraw international support/ funding for warring parties.
Grassroots mobilisation: "When the people lead, the leaders will follow." If we mobilise enough opposition to war, political leaders will be forced to bring peace.	Mobilise grassroots groups to either oppose war or to advocate positive action; use of the media; non-violent direct action campaigns; education and mobilisation effort; organising advocacy groups; dramatic or public events to raise consciousness.
Peace agreements/accords: Some form of political settlement is a prerequisite to peace – we must support a negotiation process among key parties to the violence.	Official negotiations among representatives; civil society dialogues to support negotiations; track 1½ or 2 dialogue among influential persons.
Economic action: People make personal decisions, and decision makers make policy decisions based on a system of rewards and incentives and punishment and sanctions that are essentially economic in nature. If we can change the economies associated with war making, we can bring peace.	Use of government or financial institutions to change supply and demand dynamics; control incentive and reward systems; boycotts and embargoes.
Public attitudes: War and violence are partly motivated by prejudice, misperceptions, and intolerance of difference. We can promote peace by using the media (television and radio) to change public attitudes and build greater tolerance in society.	TV and radio programmes that promote tolerance; modelling tolerant behaviour; symbolic acts of solidarity/unity; dialogue among groups in conflict, with subsequent publicity.
Transitional justice: Societies that have experienced deep trauma and social dislocation need a process for handling grievances, identifying what happened, and holding perpetrators accountable. Addressing these issues will enable people to move on to reconstruct a peaceful and prosperous society.	Truth and reconciliation commissions; criminal prosecutions and war crimes tribunals; reparations; community reconciliation processes; traditional rites and ceremonies; institutional reforms.
Community reintegration: If we enable displaced people (IDPs/refugees) to return to their homes and live in relative harmony with their neighbours, we will contribute to security and economic recovery.	Negotiation and problem solving to enable returns; intergroup dialogue; ex-combatant community engagement; processes for handling land claims; trauma healing.
Culture of peace: If we transform cultural and societal norms, values and behaviours to reject violence, support dialogue and negotiation, and address the fundamental causes of the conflict, we can develop the long-term conditions for peace.	Peace education; poverty eradication; reduction of social inequalities; promotion of human rights; ensuring gender equality; fostering democratic participation; advancing tolerance; enhancing the free flow of information and knowledge; reducing the production of and traffic in arms.

that while dialogue activities opened space for inter-ethnic interaction that might otherwise not have happened, produced some powerful effects on individuals, and led to some co-operative activities across ethnic lines, they neither strengthened community relationships nor led to collective opposition to violence.

The assumption that the changes in attitude resulting from dialogue would lead to changes in political attitudes and actions, or trickle out to influence others in the community or trickle up to influence key decision makers, proved to be wrong. In both Kosovo Albanian and Kosovo Serb communities, implicit intra-community pressures, or "rules of the game", restricted the boundaries of permissible interaction to generally non-visible business interactions and made maintenance and expansion of inter-ethnic linkages difficult.

The examples from Liberia and Kosovo illustrate just some of the common theories of change underlying policies and projects working for peace. Others are listed, along with example methods for each, in Table B.1. An initial list of these theories was derived from reviewing the case studies in *Reflecting on Peace Practice* CDA Collaborative Learning Projects (2004). The table is not intended as a check list or menu of theories. It is designed to be purely illustrative, helping to clarify the concept of theories of change and provide concrete examples. It is by no means exhaustive and the theories it lists are not mutually exclusive – one or more could underlie a single programme.

ANNEX C

Sample terms of reference for a conflict evaluation

To consider how an evaluation's terms of reference (TOR) should be drawn up, this annex takes as an imaginary example the evaluation of a peace journalism programme (schematically outlined in Figure 4.1). It is provided to give readers an idea of the type of information to include in a conflict prevention and peacebuilding TOR. It is indicative and should not be taken as a form model. A real TOR would give greater detail. Further tips on drafting TORs can be found in *Quality Standards for Development Evaluation* (OECD, 2010c).

Terms of Reference: Evaluation of the Agency's "Peace Journalism" programme in conflict area X (2000-03)

Define the purpose and use of the evaluation. Is the purpose learning or accountability? Will the evaluation be used to decide on future funding? To inform future support? To provide input to new strategy?

The purpose of this evaluation is to determine to what extent the peace journalism programme was implemented according to agency regulations (accountability), what contribution, if any, the peace journalism training course has made to reducing inter-ethnic tensions in country X and how this contribution was made (learning), and if peace journalism makes a significant contribution to long term stability (testing the theory of change). The evaluation will be published and made available to programme managers and country field staff. Management will use the findings to decide on continued funding of the programme. Findings are also of interest to the government of neighbouring country Y, which is considering a similar approach for a nationwide programme.

Describe the evaluation object and scope. What are the specific objectives of the evaluation? Is it to document achievements? Assess some or all of the activity's objectives? Will it look at implementation strategies and processes? Will the evaluation have a participatory focus? Will it look at the programme's underlying assumptions and theory of change? Which DAC evaluation criteria will be used (impact, relevance, sustainability, efficiency, or effectiveness)?

The evaluation will examine the results of the peace journalism programme from 2000-03 and its impact on peace and conflict dynamics. Specifically, it will assess whether or not peace journalism is an effective and efficient contribution to peacebuilding. The five DAC evaluation criteria will be used to assess the programme.

Describe the rationale of the evaluation: Why this evaluation at this point of time? Describe the longevity, amount of funding, and risks tied to the intervention. Are there any

specific events that have triggered the evaluation (unveiling of corruption, results that run counter to intentions in the intervention, new research being released)?

The conflict situation is worsening in country X and the public in the donor country is demanding to know how our agency has been involved in recent changes. Also, the agency is considering funding similar journalism programmes in other regions and would like to know if this is an effective strategy to pursue.

Describe the scope, timeframe, objectives and nature of the activity to be evaluated. Specify issues to be covered, budget and funds spent, the time period to be evaluated, types of activities, geographical coverage, target groups, as well as other elements of the conflict prevention and peace building intervention addressed, such as contextual issues.

The peace journalism programme involved the training of 50 journalists from eight municipal districts and four workshops for interior ministry staff. The training took place over the course of two days and were run by agency staff and local organisation partners. The total funds disbursed were EUR 500 000. The programme was meant to contribute to peace by reducing bias in reporting and making journalists more aware of the sources and dynamics of conflict in relation to their work (theory of change). Each training session involved activities led by the agency's country staff. Workshops were held. Participants included 57% women, while 30% were from the dominant religious group (70% from minority religious groups), 40% from minority ethnic group A and 60% from B.

The programme has not been reviewed. Country and programme staff provided twice yearly self-assessments showing the output and achievements of basic outcome objectives, which included the number of journalists trained. Evaluations of workshops and trainings were completed by participants. While staff felt this was a successful programme overall, recent escalations in violence have raised concerns about impact. Many participants have changed their views of the programme in light of the changing situation on the ground.

Provide directions for the approaches to be used. What method will be used in the evaluation? How should the evaluation be conducted, via what process and steps, etc.? Will there be an inception phase? What will the level of stakeholder involvement in the evaluation process be?

The evaluators will undertake a thorough conflict analysis and then draft an inception report describing how they will answer the key evaluation questions. The evaluation will include a desk review of the programme's self-evaluations and participants' evaluations, as well as spending and country reports from the agency and other donors in the region. The evaluation team will visit country X for a participatory workshop with programme staff and embassy staff, as well as to interview programme participants.

Logistical and safety concerns: Address ethical behaviour in conflict environments and provide guidance on safety and logistics.

Due to safety concerns, the evaluation team will conduct field visits in Regions 1 and 2, but not in Region 3. For Region 3, evaluators will instead meet with proxies in the capital and collect data from a recent OXFAM community study carried out in Regions 3 and 4. The visit should take place during March. Security escorts in Region 2 will be provided by the embassy of country X.

Principles: What standards and principles are to be followed. Refer the team to any relevant policy documents or agency agreements.

The evaluation should follow our agency's "Principles for Engagement in Conflict Situations" and adhere to the DAC Quality Standards. The team is also expected to adhere to our agency's "Guidelines on Gender Sensitive Development Assistance".

Management arrangements, quality control and reporting. Who will be in charge of each task and oversight? To whom will the evaluation team report? Is there a need to establish a steering mechanism for the evaluation? Who will be responsible for ensuring information sharing among team members? Who will be involved in drawing and assessing conclusions? What reports will be generated? Will they be public or confidential? Will they be published or placed on the internet? Will the reports and conclusions be checked? What quality control systems will be used?

The team will report directly to the evaluation department country programme manager Ms. X and will also work with a small reference group including X, Y, and Z who will review and comment on the inception report. The team will complete a field report which will be presented in a participatory workshop to country staff before completing the field mission. A and B will review the final report before it is accepted for publication.

Team requirements (including team make-up). Who should do the evaluation and what characteristics do they need to have? What is the desired size and composition of the team? What time commitment is involved in terms of person-hours? What types of individuals are needed for this particular evaluation in this particular context?

The team should include experts in ethnic conflict and land-based disputes with specific expertise on this conflict. There should be a team leader (40 days), at least four experts (30 days each) and up to two research assistants (30 days total). At least two members should be fluent in language A and language B and all members should be comfortable working under difficult circumstances and should have good communication skills and non-aggressive attitudes.

Budget and schedule. How will the evaluation be funded? Have there been arrangements made for security costs or other additional costs associated with working in a conflict environment? Are funds available for conflict analysis? (Bids may also be accepted and then compared to establish the appropriate funding needed.) When will the evaluation be conducted? What criteria will be applied to reports to disburse funding?

© OECD 2012

Bibliography

African Union (2006), "Policy Framework for Post-Conflict Reconstruction and Development", African Union, Addis Ababa.

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

Anderson, M.B. and L. Olson (2003), *Confronting War: Critical Lessons for Peace Practitioners,* Reflecting on Peace Practice Project, Collaborative For Development Action, Inc., Cambridge, MA.

Aoi, C., R. Thakur and C.H. de Coning (eds.) (2007), *Unintended Consequences of Peacekeeping*, United Nations University Press, New York.

Asian Development Bank (ADB), (2007), "Emergency Flood Rehabilitation Project in Tajikistan", Operations Evaluation Department, Manila.

Bamberger, M. *et al.* (2004), "Real World Evaluation in a Nutshell", a brief summary of full article "Shoestring Evaluation", *American Journal of Evaluation*, Vol. 25, No. 1, Spring.

Bamberger, M., J. Rugh and L. Mabry (2006), *Real World Evaluation Working Under Budget, Time, Data, and Political Constraints,* Sage Publications, Inc., Thousand Oaks.

Beck, T. (2006), *Evaluating Humanitarian Action: Using the OECD-DAC Criteria – An ALNAP Guide for Humanitarian Agencies*, Overseas Development Institute (ODI), London.

Bellina, S., H. Magro, and V. de Villemeur (eds.) (2009), *Democratic Governance: A New Paradigm for Development?* Hurst Publishers, London.

Bellina, S. *et al.* (2010), "L'État en quête de légitimité: sortir collectivement des situations de fragilités", Editions Charles Léopold Mayer, Paris.

Bennett, J. *et al.* (2010), "Aiding the Peace: A Multi-donor Evaluation of Support to Conflict Prevention and Peacebuilding Activities in Southern Sudan 2005-2010", Itad Ltd., UK.

Böhnke, J., J. Koehler, C. Zürcher (2007), "Assessing the Impact of Development Co-operation in North East Afghanistan: Approaches and Methods", *Evaluation Working Papers*, Bundesminsterium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ), Bonn.

Böhnke, J., J. Koehler, C. Zürcher (2010), "Assessing the Impact of Development Co-operation in North East Afghanistan 2005-2009", *Evaluation Reports 049*, *Final Report*, BMZ, Bonn.

Borton, J. and J. Eriksson (2004), "Lessons from Rwanda – Lessons for Today. Assessment of the Impact and Influence of Joint Evaluation of Emergency Assistance to Rwanda", Danish Ministry of Foreign Affairs, Danish State Information Centre, December, *www.danida.netboghandel.dk*.

Borton. J., M. Buchanan-Smith and R. Otto (2005), *Support to internally displaced persons: Learning from evaluations*, Synthesis Report on a Joint Evaluation Programme, commissioned by the Swedish International Development Agency (SIDA), the Danish and Dutch Ministries of Foreign Affairs, and the European Commission Humanitarian Aid Office (ECHO), SIDA, Stockholm.

Brabant, K. van (2001), "Understanding, Promoting and Evaluating Co-ordination: An Outline Framework" in D.S. Gordon and F.H. Toase (eds.), *Aspects of Peacekeeping*, Frank Cass, London.

Brusset, E. *et al.* (2011), *Amani Labda, Peace Maybe, Joint Evaluation of Conflict Prevention and Peace-Building in the Democratic Republic of Congo*, synthesis report, Channel Research, Belgian Ministry of Foreign Affairs, Foreign Trade and Development Co-operation, Brussels.

Canadian International Development Agency (2011), "Management Response to the Recommendations of the Multi-Donor Evaluation of Support to Conflict Prevention and Peacebuilding Activities in Southern Sudan from 2005-2010", CIDA, *www.acdi-cida.gc.ca/acdi-cida/ACDI-CIDA.nsf/eng/CAR-528105656-KJM*.

Collaborative for Development Action (CDA) Collaborative Learning Projects (2000), *Options for Aid in Conflict: Lessons from Field Experience*, CDA Collaborative Learning Project, Cambridge, MA.

CDA Collaborative Learning Projects (2004), *Reflecting on Peace Practice Project*, CDA Collaborative Learning Projects, Cambridge MA.

CDA Collaborative Learning Projects (2007), *Do No Harm* Checklist, *www.cdainc.com/cdawww/project_profile.php?pid=DNH&pname=Do%20No%20Harm*.

CDA Collaborative Learning Projects (2009), *Reflecting on Peace Practice – Participant Training Manual*, Cambridge, MA.

Chapman, N. *et al.* (2009), *Evaluation of Donor-Supported Activities in Conflict-Sensitive Development and Conflict Prevention and Peacebuilding in Sri Lanka*, Main Evaluation Report, OECD, Paris.

Chesterman, S. (2002), *Tiptoeing Through Afghanistan: The Future of UN State-Building*, International Peace Academy, New York.

Chigas, C. *et al.* (2006), *Has peacebuilding made a difference in Kosovo ? – A Study of the Effectiveness of Peacebuilding in Preventing Violence: Lessons Learned from the March 2004 Riots in Kosovo*, CDA Collaborative Learning Projects and Care International, Pristina.

Chr. Michelsen Institute, Copenhagen Development Consulting, and the German Association of Development Consultants (2005), *Joint Evaluation: Humanitarian and Reconstruction Assistance to Afghanistan, 2001-05 from Denmark, Ireland, the Netherlands, Sweden and the United Kingdom,* Chr. Michelsen Institute, October.

Church, C. and M. Rogers (2006), "Designing for Results: Integrating Monitoring and Evaluation in Conflict Transformation Programs, Search for Common Ground", Washington, DC.

Church, C. and J. Shouldice (2003), "The Evaluation of Conflict Resolution Interventions, Part II: Emerging Practice and Theory", INCORE, London.

Coffey International Development, "Helmand Monitoring and Evaluation Programme", *www./uk.coffey.com/our-projects/helmand-monitoring-and-evaluation-programme*.

Collier, P. *et al.* (2003), "Breaking the Conflict Trap: Civil Wars and Development Policy", World Bank and Oxford University Press, Washington, DC.

Conflict Sensitivity Consortium (2004), "An Introduction to conflict-sensitive approaches to development, humanitarian assistance and peace building: tools for peace and conflict impact assessment tools for peace and conflict impact assessment", Resource Pack, Chapter 2, Conflict Sensitivity Consortium.

Coning, C.H. de (2007), "Civil-Military Co-ordination and UN Peacebuilding Operations" in H. Langholtz, B. Kondoch, A. Wells (eds.), International Peacekeeping: The Yearbook of International Peace Operations, Vol. 11, Koninklijke Brill N.V., Brussels.

Cosgrave, J., B. Ramlingam, T. Beck (2009), "Real Time evaluations of humanitarian action", ALNAP Guide, pilot version, Overseas Development Institute, London, *www.alnap.org/pool/files/rteguide.pdf*.

CSIS (2002), "Post-Conflict Reconstruction: Task Framework, Center for Strategic and International Studies (CSIS) and the Association of the United States Army (AUSA)", Washington, DC.

Cutillo, A. (2006), "International Assistance to Countries Emerging from Conflict: A Review of Fifteen Years of Interventions and the Future of Peacebuilding, International Peace Academy (IPA)", New York.

Dahrendorf, N. (2003), "A Review of Peace Operations: A Case for Change: East Timor Study, Conflict, Security and Development Group", King's College, London.

Danish International Development Assistance (Danida) (2011), "Open Fun Football Schools Programme", Evaluation Study, Ministry of Foreign Affairs of Denmark, Copenhagen.

Development Evaluation Resource Centre (DEReC), OECD DAC Network on Development Evaluation, *www.oecd.org/dac/evaluationnetwork/derec*.

DevInfo (2012), "DevInfo In Action: Occupied Palestinian Territory: Using DevInfo to Provide Data for Decision-Making", *DevInfo Newsletter*, 19 January 2012, *www.devinfo.org/devinfo_in_action.html*.

Dobbins, J.E. *et al.* (2005), "The UN's Role in Nation-Building: From Congo to Iraq, Rand Corporation", Washington, DC.

Donini, A. (2002), "The Policies of Mercy: UN Co-ordination in Afghanistan, Mozambique and Rwanda", *Occasional Paper #22*, Thomas J. Watson Jr. Institute for International Studies, Brown University, Providence, RI.

Doyle, M.W. and N. Sambanis (2000), *International Peacebuilding: A Theoretical and Quantitative Analysis*, World Bank, Washington, DC.

Duffield, M., S. Lautze, B. Jones (1998), *Strategic Humanitarian Co-ordination in the Great Lakes Region 1996-1997*, United Nations Office for the Co-ordination of Humanitarian Affairs (OCHA), New York.

Eriksson, J. *et al.* (1996), *The International Response to Conflict and Genocide: Lessons from the Rwanda Experience – Synthesis Report, Joint Evaluation of Emergency Assistance to Rwanda,* DANIDA, Copenhagen.

European Commission (2011), "Thematic Evaluation of European Commission Support to Conflict Prevention and Peace-building", *Final Report*, Commission of the European Union, Brussels. *www.oecd.org/dataoecd/28/22/49446993.pdf*.

Fafo Institute for Applied International Studies (2006), *Evaluating Conflict Prevention and Peacebuilding Activities*, report prepared for the OECD DAC Network on Development Evaluation and the Network on Conflict, Peace and Development Co-operation, Fafo, Oslo, unpublished.

Feinstein, Osvaldo (2002), p. 433: "Use of Evaluations and the Evaluation of their Use" in *Evaluation,* Vol. 8 (No. 4), pp. 433-439, *www.evi.sagepub.com/cgi/content/abstract/8/4/433*.

Ficatier, Y. and the Groupe de recherches et d'échanges technologiques (Technology Research and Exchange Group) (GRET) (2010), *Operating in Fragile States: Lessons from Experience*, Summary Notes Series, No. 9, Evaluation and Capitalisation Unit, Agence française de développement (French Development Agency) (AFD), October.

French Ministry of Foreign and European Affairs (2006), "Governance Strategy for French Development Assistance", Directorate General for International Co-operation and Development (DGCID), French Ministry of Foreign and European Affairs, Paris, December.

French Ministry of Foreign and European Affairs (2008), France, *La Politique de la France dans les pays en situation de fragilité: Evaluation des interventions de la coopération et des ONG françaises dans des contextes de crise et de fragilité institutionnelle (2000-2007)*, Evaluation No. 103, DGCID, French Ministry of Foreign and European Affairs, Paris, June.

Gaarder, M. (2011), "3iE's Approach to Policy Influence", PowerPoint Presentation to the 3iE Member's Meeting, 3iE, May 2011, *www.3ieimpact.org*.

Glouberman, S. and B. Zimmerman (2002), "Complicated and Complex Systems: What Would Successful Reform of Medicare Look Like?", *Discussion Paper No. 8*, Commission on the Future of Health Care in Canada, *www.hc-sc.gc.ca/english/pdf/romanow/pdfs/8_Glouberman_E.pdf*.

Good Humanitarian Donorship (GHD) (2003), *Principles and Good Practice of GHD*, *www.goodhuman itariandonorship.org*.

Goor, L. van de and S. Verstegen (2000), "Conflict Prognosis: A Conflict and Policy Assessment Framework", Clingendael Institute, The Hague.

Humphreys, M., R. Sanchez de la Sierra, P. van der Windt (2012), "Social and Economic Impacts of Tuungane: Final Report on the Effects of a Community Driven Reconstruction Program in Eastern Democratic Republic of Congo", Columbia University, New York.

IASC (2004), *Civil-Military Relations in Complex Emergencies*, UN Inter-Agency Standing Committee (IASC), United Nations, New York, June.

International Alert (2001), "Gender and Peace Support Operations: Opportunities & Challenges to Improve Practice", Gender and Peacebuilding Programme, International Alert, London. *www.international-alert.org/pdfs/pso_briefing_2001.PDF*.

International Alert (2004), *Oil and the Search for Peace in the South Caucasus: The Baku-Tbilisi-Ceyhan (BTC) oil pipeline*, International Alert, London.

International Alert (2005), "Conflict-Sensitive Programme Management in International Co-operation", Conflict Sensitivity Consortium, *www.conflictsensitivity.org*.

International Alert (2007a), "Conflict Analysis", in *Conflict-sensitive Approaches to Development, Humanitarian Assistance and Peace Building: Tools for Peace and Conflict Impact Assessment*, Conflict Sensitivity Consortium, Chapter 2, *www.conflictsensitivity.org*.

International Alert (2007b), *Acting on Commitments: How EU Strategies and Programming Can Better Prevent Violent Conflict*, International Alert, Saferworld, and the European Peacebuilding Liaison Office.

Johnston, N. (2005), "Evaluation of DFID Development Assistance: Gender Equality and Women's Empowerment – Phase II Thematic Evaluation: Conflict and Post Conflict Reconstruction", *Working Paper* 5, DFID, London, *www.dfid.gov.uk/Documents/publications1/evaluation/wp8.pdf*.

Jones, R., P. George, C. Stanley (2001), "Canada-Japan Joint Peacebuilding Learning Project Guatemala Mission Report", posted on OECD website at *www.oecd.org/dataoecd/4/9/35080245.pdf*.

Kennedy-Chouane, M.G. (2011), "Improving conflict prevention and peacebuilding assistance through evaluation", *OECD Journal: General Papers*, Volume 2010/1, OECD, Paris, 25 May, pp. 99-107.

Lederach, J.P. (2002), *Building Peace: Sustainable Reconciliation in Divided Societies*, USIP Press, Washington, DC.

Leeuw, F. and J. Vaessen (2009), *Impact Evaluation and Development: NONIE Guidance on Impact Evaluation*, commissioned by the Network of Networks for Impact Evaluation (NONIE), World Bank, Washington, DC.

Loode, S. (2010), "*Working in complex communities: what peacebuilding can learn from complex systems science*", paper presented at the 2010 IPRA Global Conference 'Communicating Peace', University of Sydney, Australia, 6-10 July 2010.

Lund, M. (2004), *Preventing Violent Conflicts: Conflict-Sensitive Development in the 21st Century*, Commissioned by the Conflict Prevention and Reconstruction Unit, Social Development Department, World Bank, Management Systems International, Inc., Washington, DC.

Menkhaus, K. (2006), "Conflict Prevention and Human Security: Issues and Challenges" in R. Picciotto and R. Weaving (eds.), *Security and Development: Investing in Peace and Prosperity*, Routledge, New York.

Minear, L. and U. Chellia (1992), *UN Co-ordination of the International Humanitarian Response to the Gulf Crisis*, Thomas J. Watson Institute for International Studies, Providence, RI.

Ministry of Foreign Affairs of Japan and the United Nations Development Programme (UNDP) (2002), *Post-conflict Assistance of the Government of Japan through UNDP in Kosovo and East Timor*, UNDP Evaluation Office, New York.

New Partnership for Africa's Development (NEPAD) (2005), *African Post-Conflict Reconstruction Policy Framework*, NEPAD, Midrand.

Norwegian Agency for Development Co-operation (Norad) (2006), "Proceedings of the Workshop on Developing a Guidance for Evaluating Conflict Prevention and Peacebuilding Activities", 30 November-1 December, Oslo.

Norad (2009), *Evaluation of Norwegian Support to Peacebuilding in Haiti 1998-2008*, Evaluation Report 5/2009, Evaluation Department, Norad, Oslo.

Norad (2011), *Pawns of Peace – Evaluation of Norwegian Peace Efforts in Sri Lanka 1997-2009*, Evaluation Report 5/2011, Evaluation Department, Norad, Oslo.

OECD (1991), *DAC Principles for Evaluation of Development Assistance*, Development Assistance Committee (DAC), OECD, Paris.

OECD (1997), *Helping Prevent Violent Conflict*, The DAC Guidelines, DAC, OECD, Paris.

OECD (1999), "Guidance for Evaluating Humanitarian Assistance in Complex Emergencies", DAC Network on Development Evaluation OECD, Paris.

OECD (2000), "Effective Practices in Conducting a Joint Multi-donor Evaluation", *Evaluation and Effectiveness*, No. 4, Development Assistance Committee (DAC), OECD, Paris.

OECD (2001a), *Evaluation Feedback for Effective Learning and Accountability*, DAC Network on Development Evaluation, OECD, Paris.

OECD (2001b), *Helping Prevent Violent Conflict*, The DAC Guidelines, supplement to the 1998 edition of *Helping Prevent Violent Conflict*, DAC, OECD, Paris.

OECD (2001c), *Poverty Reduction,* The DAC Guidelines, DAC, OECD, Paris.

OECD (2002), *Glossary of Key Terms in Evaluation and Results Based Management*, DAC Network on Development Evaluation, OECD, Paris.

OECD (2005), *The Paris Declaration on Aid Effectiveness*, declaration adopted at DAC meeting in March 2005, OECD, Paris, *www.oecd.org/dataoecd/11/41/34428351.pdf*.

OECD (2006), *Guidance for Managing Joint Evaluations*, DAC Evaluation Series, DAC, OECD, Paris.

OECD (2007a), "Principles for Good International Engagement in Fragile States and Situations", OECD, Paris, *www.oecd.org/dataoecd/61/45/38368714.pdf*.

OECD (2007b), "Encouraging effective evaluation of conflict prevention and peacebuilding activities – Towards DAC Guidance", offprint the *OECD Journal on Development*, Vol. 8, No. 3, OECD, Paris.

OECD DAC (2007c), *Manual on Security System Reform: Supporting Security and Justice*, DAC Network on Conflict, Peace and Development Co-operation, OECD, Paris.

OECD (2008a), "State Building in Situations of Fragility – Initial Findings", Fragile States Group (FSG), DAC, OECD, Paris, August.

OECD (2008b), *Guidance on Evaluating Conflict Prevention and Peacebuilding Activities – Working Draft for Application Period*, Development Assistance Committee (DAC), OECD, Paris.

OECD (2010a), "Accelerating Progress towards the Millennium Development Goals in Countries affected by Conflict, Violence and Insecurity", note prepared for DAC Senior Level Meeting on 14-15 April, International Network on Conflict and Fragility (INCAF), DAC, OECD, Paris, *www.violenceanddevelopment.org/docs/inaf_accelerating_progress_towards_mdgs.pdf*

OECD (2010b), *Evaluating Development Co-operation: Summary of Key Norms and Standards*, 2nd Edition, OECD DAC Network on Development Evaluation, OECD, Paris, *www.oecd.org/dataoecd/12/56/41612905.pdf*.

OECD (2010c), *Quality Standards for Development Evaluation*, DAC Guidelines and Reference Series, Development Assistance Committee (DAC), OECD, Paris.

OECD (2010d) *Do No Harm: International Support for Statebuilding,* Conflict and Fragility, OECD, Paris, *www.oecd.org/dataoecd/8/32/44409926.pdf*.

OECD (2011a), *Evaluating Development Co-operation: Summary of Key Norms and Standards,* OECD, Paris, *www.oecd/org/dac/evaluation*.

OECD (2011b), *Supporting Statebuilding in Situations of Conflict and Fragility: Policy Guidance*, DAC Guidelines and Reference Series, DAC, OECD, Paris, *www.oecd.org/document/12/0,3746,en_2649_33693550_46623180_1_1_1_1,00.html*.

OECD (2011c), "The New Deal for International Engagement in Fragile States", agreement endorsed at the Fourth High Level Forum on Aid Effectiveness meeting in Busan from 29 November to 1 December.

OECD (2011d), *International Engagement in Fragile States: Can't we do better?*, OECD, Paris, *www.oecd.org/fsprinciples*.

OECD (2011e), *Managing Risks in Fragile and Transitional Contexts: The Price of Success?,* OECD, Paris.

OECD (2011f), "Ensuring Fragile States Are Not Left Behind – 2011 Factsheet on Resource Flows in Fragile States", INCAF, DAC, OECD, Paris, *www.oecd.org/dataoecd/5/11/49108935.pdf*.

OECD (2012), *International Support to Post-Conflict Transition: Rethinking Policy, Changing Practice*, DAC Guidelines and Reference Series, OECD Publishing, Paris, *DOI: 10.1787/9789264168336-en*.

OECD and CDA Collaborative Learning Projects (2007), *Encouraging Effective Evaluation of Conflict Prevention and Peacebuilding Activities: Towards DAC Guidance*, consultative report prepared for the OECD DAC Network on Development Evaluation and the DAC Network on Conflict, Peace and Development Co-operation, OECD, Paris.

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

Paffenholz, T. (2011), "Evaluation Reports: The German Civil Peace Service, Synthesis Report, Volume I: Main Report", BMZ, Bonn, *www.oecd.org/dataoecd/55/47/49139167.pdf*.

Picciotto, R. and R. Weaving (eds.) (2006), *Security and development: investing in peace and prosperity,* Routledge, London.

Polastro, R. *et al*. (2011), "IASC Evaluation of the Humanitarian Response in South Central Somalia 2005-2010," DARA, Madrid, *www.oecd.org/dataoecd/9/29/49335639.pdf*.

Pugh, M. and N. Cooper (2004), *War Economies in a Regional Context*, Lynne Rienner, London.

Reindorp, N. and P. Wiles (2001), *Humanitarian Co-ordination: Lessons from Recent Field Experience*, study commissioned by OCHA, Overseas Development Institute (ODI), London.

Scriven, M. (1991), *Evaluation Thesaurus*, Fourth Edition, Sage Publications Inc., Newbury Park, CA.

Search for Common Ground: The Learning Portal for Design, Monitoring and Evaluation for Peacebuilding, *www./dmeforpeace.org*.

Swedish International Development Agency (SIDA) (2003), *Reflection on Experiences of Evaluating Gender Equality*, SIDA Studies in Evaluation, 03/01, Department for Evaluation and Internal Audit, Stockholm.

SIDA (2004), *Conflict Sensitive Development Co-operation: How to Conduct a Conflict Analysis,* Swedish Development Agency (SIDA), Stockholm.

Smith, D. (2004), *Towards a Strategic Framework for Peacebuilding: Getting Their Act Together*, Overview Report of the Joint Utstein Study on Peacebuilding, Ministry of Foreign Affairs, Oslo.

Stedman, S.J. (2005), "A Peacebuilding Commission as An Instrument of Conflict Prevention" in A. Melbourn (ed.), *Development, Security and Conflict Prevention: Security as a Millennium Goal,* Anna Lindh Programme on Conflict Prevention, Madariaga European Foundation, Stockholm.

Stedman, S.J., E. Cousens and D. Rothchild (eds.) (2002), *Ending Civil Wars: The Implementation of Peace Agreements*, Lynne Rienner, Boulder, CO.

Stockton, N. (2002), *Strategic Co-ordination in Afghanistan*, Afghanistan Research and Evaluation Unit (AREU), Kabul.

Strand, A. (2003), *Who's Helping Who? NGO Co-ordination of Humanitarian Assistance*, Doctoral Dissertation, Department of Politics, University of York, York.

Tschirgi, N. (2006), *The Security-Development Nexus: Research Findings and Policy Implications*, International Peace Academy (IPA), New York.

United Nations (UN) (1992), *An Agenda for Peace: Preventive Diplomacy, Peacemaking and Peace-keeping*, United Nations, New York.

UN (2000), *We the Peoples: The Role of the United Nations in the 21st Century*, Report of the Secretary-General to the 54th Session of the General Assembly, A/54/2000, United Nations, New York, 27 March.

UN (2001), *Prevention of Armed Conflict*, Report to the Secretary General, United Nations, New York.

UN (2007), Decision of the Secretary-General's Policy Committee, United Nations, New York, May.

UN (2010), *Monitoring Peace Consolidation- United Nations Practitioners' Guide to Benchmarking*, United Nations, New York.

United Nations Development Programme (2002), "Gender Approaches in Conflict and Post-conflict Situations", *www.undp.org/women/docs/gendermanualfinalBCPR.pdf*.

United Nations Security Council (UNSC) (2000), "Resolution 1325", adopted by the Security Council at its 4213th meeting, 31 October 2000, United Nations, New York, *www.un.org/events/res_1325e.pdf*.

United States Agency for International Development (USAID) and Management Systems International (2006), "Post conflict monitoring and evaluation", USAID, Washington, DC.

United States Agency for International Development (USAID) (2007a), *Women & Conflict – An Introductory Guide for Programming, USAID*, Washington, DC.

USAID (2007b), "Conflict Management", *www.usaid.gov/our_work/cross-cutting_programs/conflict* .

Uvin, P. (1999a), "Incentives and Disincentives for Influencing Conflict Situations", *DAC Journal*, Vol. 2, No. 3, OECD DAC, Paris, *www.oecd.org/dataoecd/31/24/1843044.pdf*.

Uvin, P. (1999b), "The Influence of Aid in Situations of Violent Conflict: A Synthesis and Commentary on the Lessons Learned from Case Studies on the Limits and Scope for the Use of Development Assistance", Informal Task Force on Conflict, Peace, and Development Co-operation, Development Assistance Committee, OECD, Paris.

Uvin, P. (2002), "The Development/Peacebuilding Nexus: A Typology and History of Changing Paradigms", *Journal of Peacebuilding & Development*, Vol. 1, No. 1.

Vogel, I. (2012), *Review of the use of "Theory of Change" in international development*, draft review report for the UK Department for International Development, London.

Weiss, C. (1995), "Nothing as Practical as Good Theory: Exploring Theory-based Evaluation for Comprehensive Community Initiatives for Children and Families", in J. Connell *et al.* (eds.), *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, Aspen Institute, Washington, DC.

Wood, B. (2003), *Development Dimensions of Conflict Prevention and Peacebuilding*, UN Development Programme (UNDP), New York.

World Bank (2005a), *Towards a Conflict-Sensitive Poverty Reduction Strategy: Lessons from a Retrospective* Analysis, World Bank, Washington, DC.

World Bank (2005b), "Conducting Quality Impact Evaluations under Budget, Time and Data Constraints", Independent Evaluation Group, World Bank, Washington, DC.

World Bank (2011), *Conflict, Security, and Development,* World Development Report 2011, World Bank, Washington, DC.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where governments work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Commission takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

DEVELOPMENT ASSISTANCE COMMITTEE (DAC)

In order to achieve its aims, the OECD has set up a number of specialised committees. One of these is the Development Assistance Committee (DAC), whose mandate is to promote development co-operation and other policies so as to contribute to sustainable development – including pro-poor economic growth, poverty reduction and the improvement of living standards in developing countries – and to a future in which no country will depend on aid. To this end, the DAC has grouped the world's main donors, defining and monitoring global standards in key areas of development.

The members of the DAC are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Korea, Luxembourg, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, the United States and the European Union.

The OECD Development Assistance Committee develops guidelines and reference documents, published in the DAC Guidelines and Reference Series, to inform and assist members in the conduct of their development co-operation programmes.

OECD PUBLISHING, 2, rue André-Pascal, 75775 PARIS CEDEX 16
(43 2012 15 1 P) ISBN 978-92-64-10679-6 – No. 60259 2012

DAC Guidelines and Reference Series

Evaluating Peacebuilding Activities in Settings of Conflict and Fragility

IMPROVING LEARNING FOR RESULTS

Contents

Chapter 1. Conceptual background and the need for improved approaches in situations of conflict and fragility

Chapter 2. Addressing challenges of evaluation in situations of conflict and fragility

Chapter 3. Preparing an evaluation in situations of conflict and fragility

Chapter 4. Conducting an evaluation in situations of conflict and fragility

Please cite this publication as:

OECD (2012), *Evaluating Peacebuilding Activities in Settings of Conflict and Fragility: Improving Learning for Results,* DAC Guidelines and Reference Series, OECD Publishing.
http://dx.doi.org/10.1787/9789264106802-en

This work is published on the *OECD iLibrary,* which gathers all OECD books, periodicals and statistical databases. Visit ***www.oecd-ilibrary.org***, and do not hesitate to contact us for more information.

OECD*publishing*
www.oecd.org/publishing

ISBN 978-92-64-10679-6
43 2012 15 1 P