Chair's summary

Evaluating conflict prevention and peacebuilding activities:
Lessons from the application phase

On 16 and 17 February 2011, around 70 experts and policy makers from a wide range of OECD member states and international development and research institutions met in Oslo to review the lessons that have emerged from piloting the draft *DAC Guidance on Evaluating Conflict Prevention and Peacebuilding Activities*. This draft guidance has been used for major joint evaluations of multi-year donor policies and programmes in the area of Conflict Prevention and Peace Building (CPPB) in three fragile states: Southern Sudan, Sri Lanka and the D.R. Congo and was also used in single donor evaluations in Haiti and Afghanistan and in Germany to evaluate a global fund for peace service. Evaluation reports and further background on the draft guidance are available here. The workshop aimed to review both policy and evaluation lessons emerging from these evaluations and to provide input and suggestions for finalization of the Guidance.

Overall conclusions

As emphasised in the opening address of Norad Director General Villa Kulild, the guidance was intended to help fill a "strategic gap" in these fields, at a time when the international understanding of conflict and fragility was relatively limited and there was little agreement on how donors should work on conflict prevention and peacebuilding. The strategic gap in peacebuilding is reflected both in terms of weak programme design and in a dearth of rigorous evaluations. The testing phase thus aimed at evaluating a range of conflict prevention and peacebuilding programmes – producing policy insights as well as feedback on the usefulness of the draft guidance. The wealth of knowledge shared during the event demonstrated that, while many challenges remain, quality evaluation is possible.

The guidance also aimed to bridge a gap between two communities: introducing conflict and fragility programme and policy communities to the concepts and use of evaluation and introducing peace and conflict specificities to evaluation experts. Participants reported that interest in and use of the guidance was high, not just among INCAF and EVALNET members, but also from educational institutions, NGOs, researchers in developing countries and the military. The document has been widely disseminated (in both English and French) during the three year test phase.

Participants re-emphasized that fragile and conflict affected states constitute highly complex and fluid environments, which has important implications for the conceptualization, design, implementation and evaluation of CPPB activities and require different approaches to understanding the role of aid in change processes. Participants welcomed the work that has been conducted over the past couple of years to better understand these issues, including the OECD's Statebuilding Guidance, work in the area of transition financing and risk management and the ongoing process of the International Dialogue on Peacebuilding and Statebuilding. Such work has progressed greatly, strengthening the analytical framework for conflict prevention, peacebuilding and statebuilding activities – and their evaluation.

Conflict prevention and peacebuilding findings and policy lessons

The evaluation findings presented demonstrated that there is plenty of scope for improvement of donor activities in these areas. While the performance of specific programmes varied, a number of common themes emerged. In particular international actors working on conflict and fragility need to better understand and address: a) linkages between national, capital focused peace processes and regional/local

activities; b) engagement based on technical vs. political mandates; c) focus on achievement of socio-economic development objectives or specific activities aimed at peacebuilding and addressing conflict drivers; d) the need for longer term engagement, but also for flexibility in funding mechanisms, sector priority and geographic location; e) the challenges around working with / around governments in situations of low capacity and weak legitimacy; f) understanding the specificity of context, conflict and change processes and the (limited) role donors can play; g) the need for better programming and evaluative thinking from the outset of programme design.

Evaluation teams highlighted several policy lessons for international engagement that have emerged from the evaluations. In particular, they suggest that donors should:

Understand the context better: at home and abroad

Ground interventions in solid conflict analysis. Good conflict analysis needs to include an analysis of the political economy, context and stakeholders. The use of conflict analysis should be flexible and strategic and analysis should therefore be monitored and updated on a regular basis. It should focus on drivers of conflict rather than root causes. This requires significant resources, but is vital to the relevance and effectiveness of engagements. Numerous tools for analysis are already available, of which the guidance provides an overview. Further guidance on the use of conflict analysis in the evaluation process is needed (see evaluative lessons below).

Recognize the politics of aid and its implications for local ownership. External support is not politically neutral; it provides resources and often tries to influence interest. It does not do so in a vacuum. Especially in fragile and conflict-affected states the stakes are high because of conflict legacies and limited resource availability. To stimulate gradual change, donor engagements need to contain incentives for elites if they are to generate ownership beyond the marginalized groups of a society.

Recognize the influence of domestic donor politics on aid effectiveness, in particular in the case of *Whole of Government* efforts. Donors should take care that such engagements are based on shared context analysis, objectives, strategies and resources as much as possible. Internally, they need to recognize the political realities in which they pursue their objectives and deal with the trade-offs involved.

Develop more strategic frameworks for engagement

Ensure peace- and statebuilding engagements work to find balance between state responsiveness and societal needs and expectations. This needs to be grounded in a good understanding of what is locally possibly and desirable. Donors must beware of supporting statebuilding defined solely in Weberian terms or based on unrealistic expectations of what is attainable in a particular context. Among others, the OECD's and Norad's recent work on statebuilding, political legitimacy and political settlements offers conceptual and practical entry points to avoid doing so.

Identify and target the sectors and issues that matter most from a peacebuilding and statebuilding perspective. All of the evaluations showed that donors lacked (to a greater or lesser extent) strategic prioritisation in terms of aiming programmes and funding at addressing core peace drivers /conflict mitigating factors. While these are often the most difficult areas to work on in conflict settings, prioritising broad socio-economic development or humanitarian aid, does little or nothing to improve chances for peace – and can therefore ultimately undermine progress made in other areas. Too much funding is targeting standard socio-economic sectors (more than 80% in the case of South Sudan), while only limited attention is given to specific CPPB sectors, such as security and justice in South Sudan or the mining sector in the DRC. The evaluations show that the focus on basic services tends to be based on the flawed assumption that all development activities contribute to peacebuilding. This assumption was not supported

by results from the evaluations – in fact in some cases an over-emphasis on service delivery was seen to undermine peacebuilding prospects by diverting attention and resources away from critical elements such as land reform, justice and security sector or state capacity development.

Ensure that short-term activities support long-term strategic priorities. Donors need to base their interventions not only on the need to support short term stability but on a broader understanding of their likely impacts on longer term statebuilding, grounded in the country's possible transformational trajectory. Individual initiatives and projects might have a positive short-term effect on conflict dynamics, but if they do not benefit from such linkage, they will have limited effect in relation to long-term prospects.

Improve the logic and relevance of interventions: The evaluations highlighted that donors often apply concepts and terms without due consideration to their accurate meaning. As highlighted in the case of South Sudan, the inaccurate assumption that all development aid contributes to CPPB resulted in the vast majority of aid being spent on standard socio-economic activities rather than to target specific activities that could support sustainable peace. Similarly, the evaluations suggested that CPPB aspects are sometimes included in regular development projects without due attention to or assumptions about how a specific activity might contribute to peace. In highly volatile contexts, such mislabelling of activities can have perverse outcomes if done without closely examining and possibly changing the focus and design to target specific CPPB objectives.

Improve programming

Allow for greater flexibility of engagement in fragile states: CPPB engagement is often too large and ambitious for the context and the instruments used for implementation are not tailored to the challenges faced in fragile states. For instance, the evaluation in southern Sudan found that donors were locked-in by 3-5 year spending plans that reduced their ability to shift geographical or sector focus in response to the evolving conflict situation. Similarly, in the case of Sri Lanka, many donors did not have the strategic and financial flexibility needed to change approaches and implementation channels once the parties pulled out of the ceasefire. Greater flexibility is required to effectively address changing priorities and to enable rapid reactions to emerging opportunities to support CPPB. This can be achieved through setting more realistic objectives for interventions (using conflict analysis and scenario planning for instance, but also through assessments of absorptive capacity etc), through designing plans and instruments that can respond flexibly (the right mix of instruments is needed to meet different objectives, while avoiding the fragmentation of instruments highlighted in the case of South Sudan), and through deploying more and better qualified staff in the field over prolonged periods of time. USAID, for instance, managed to deploy more staff in South-Sudan over a longer period, with the effect that their activity portfolio performed better than many other donors and most of the pooled funds.

Strengthen programming: Challenges remain on the "preconditions" for evaluation. Far too many development programmes lack basic good programming elements (such as clearly defined goals, objectives and strategies). Evaluation can contribute, but is not a replacement for good programme design and monitoring. There is a pressing need to improve conflict and fragility programming and to incorporate latest thinking on how donors can engage in these settings into strategies in development agencies and beyond. The outcomes of the testing phase can contribute to demonstrating the usefulness of evaluation and be used to advocate for more evaluation and more strategic approaches.

Key evaluation lessons

In addition to the broad policy messages, a number of key lessons on evaluation and the usefulness of the guidance emerged from the discussions.

Overall conclusions: The Guidance is a useful introduction to evaluation in this field, especially for practitioners with limited experience of evaluation. Particularly useful is the overview of key steps and the development of the evaluation questions, the introduction of conflict analysis and theories of change (see further discussion below) as well as the conceptual clarification provided by the document (e.g. on conflict sensitivity). The guidance provides a useful overall framework of references and sources and helpfully describes how to tackle specific matters (e.g. timing and logistic issues). However, the Guidance does not tackle the "how to" questions: conflict analysis (testing theories of change, baselines, relationship between conflict analysis and evaluation criteria, etc.) and several teams suggested that more concrete operational guidance was needed. More guidance would be useful to tackle the specific complexity of multi-country thematic evaluations and to evaluate mainstreaming of conflict prevention and peacebuilding across donor portfolios.

The politics of evaluation: it was widely recognized that evaluation – like all aid-related work – is not simply a technocratic process, but also highly political. Evaluators often found it difficult to maintain a safe, credible "evaluation space" in contexts where donors and partner governments were wary and evaluation questions highly sensitive (even potentially dangerous in some settings). Defending the integrity of evaluation findings in highly politicized and even dangerous settings can pose problems for evaluation teams, particularly where there were worries that evaluation findings   may be misused by different parties to the conflict. This shared challenge could use more in-depth and concrete treatment in the guidance. Confidentiality of data and the lack of support from political actors for frank or challenging evaluations posed serious problems during some of the pilots. Encouragingly, widespread interest in the evaluation findings (once completed) was also reported – including from Ministries of Foreign Affairs, military actors and political leaders.

Theories of Change: One of the core evaluation elements introduced in the guidance is the "theory of change". A theory of change describes the assumed (or hoped) causal relationship underlying the conflict prevention or peacebuilding activity or policy and its (intended) effects on larger peacemaking goals. The material on theories of change was widely viewed as a key contribution of the guidance.

However, most donors do not articulate a clear, evidence-based basis for intervening and it can be challenging for evaluation teams to identify underling theories of change. Evaluation in these contexts should take into account that these are the product of 'political' and institutional processes, with different interested parties, and may be deliberately left open to allow a broad interpretation. For instance how *intentionally* unspoken objectives / vague strategies should be dealt with remains a question. For instance, in Southern Sudan, donors were hesitant to emphasise statebuilding of the future South Sudan as this would have been seen as presuming the outcome of the referendum.

Some found theories of change better for project level analysis, for instance in the single donor evaluation in Afghanistan. There, it was useful to the team in that it helped them focus and deepen their analysis of plans, concepts and the military doctrine central to all military engagement in Afghanistan (COIN).

Discussions at the seminar revolved around whether it is possible (and not too demanding) to elicit individual theories of change across multiple donors based on an analysis of their policies. Moreover it was discussed whether it is useful in the case of a multi-donor evaluation to try to identify a common theory of change, as was done in the Sudan evaluation. Applying a theory of change approach at a cross-country, policy or strategic level was seen as problematic as this level of analysis would inevitably cover a multitude of different theories of change.

In conclusion, eliciting theories of change and the assumptions on which they are based can be a time consuming exercise. The revised guidance should more clearly address how to incorporate the theories of change analysis, best practices and why and when it should be used/applied.

Conflict analysis: More discussion of the use of Conflict Analysis in evaluations, challenges encountered, questions of scope and design, and issues around using an existing CA versus putting together a new one should be included in the Guidance. Discussion of how to move from Conflict Analysis (more abstract) to the Evaluation Matrix questions (very definitive) could also be referred to.

Applying lessons across conflicts: Tensions remain on the generalizability (external validity) of evaluation findings. The evaluation findings all pointed to the need for "deep" contextual understanding and continuous analysis. The discussions highlighted the importance of grounding interventions and policy in conflict realities and having a flexible, profound understanding of the context, change processes and the role of donors. How can this need for specificity / groundedness and profoundly contextualized findings and recommendations be translated into strategic level lessons applicable across countries, remains challenging. The guidance did not provide enough information on its use for multi-country or thematic evaluations and how conflict analysis and other tools could be used for analysis in such evaluations.

Samples: There was a discussion around the sampling of projects and programmes, and sample size and composition. The challenges around getting suitable coverage or representative samples is not particular to conflict and peace evaluations; such issues arise in any qualitative exercise. There was some question about whether the sample should be selected before the evaluation starts or by the evaluation team during the inception phase and the self-selection of projects by involved donors in multi-donor evaluations. The methodological implications of how a sample is selected may need to be discussed. Examples from the pilot evaluations could be included for discussion and illustration in a revised guidance.

Conflict sensitivity: A number of participants stated that more guidance would have been useful on how to evaluate the conflict sensitivity of development and peace interventions. The importance of evaluating the conflict sensitivity of interventions working 'on conflict' and also, more broadly, other types of interventions working 'in conflict', was pointed out.

Utstein (conflict prevention and peacebuilding) categories: The guidance proposes four broad categories of conflict prevention and peacebuilding work as a means of framing the field. Some participants stated that the Utstein palette provides a useful way of *describing* conflict prevention and peacebuilding (and therefore defining the scope of the evaluation); and for emphasizing that the breadth of CPPB activities goes beyond socio-economic & mediation interventions. However, it should be made clear in the Guidance that it does **not** provide an analytical framework, as there is no underlying conceptual basis (e.g. between the pillars, sequencing, prioritisation, etc.). The use and limitations of the CPPB categories should be better highlighted in a revised Guidance and should incorporate recent thinking on the links between good governance, statebuilding and peacebuilding.

Evaluation criteria: Detailed descriptions of the core DAC evaluation criteria were found to be useful, and it was underlined in the Afghanistan evaluation that the criteria *relevance*, *sustainability*, *coverage* and *policy coherence* are all well developed in the guidance. The use of the *impact* criteria in conflict prevention and peacebuilding evaluations could be problematised and expanded upon in the guidance. The criteria *linkages* and *consistency with values* are less well developed and some thought these were unnecessary additions. The German civil peace service evaluation and other meeting participants underlined the usefulness of the *policy coherence* or 3c's criteria (*coherence*, *coverage* and *coordination*) due to the interlinked nature of the field. On the other hand, some participants raised concerns and scepticisms on the necessity of the additional criteria and suggested that these additional criteria can be incorporated into the basic five DAC criteria. Discussion did not reach a consensus, and it remained to be resolved.

<u>Involvement of stakeholders:</u> The evaluation management teams in several countries dealt with tensions between the need for conflict sensitivity and the involvement of country stakeholders in the evaluation. When the government is party to an ongoing conflict, its involvement may compromise the credibility and independence of the evaluation process. Creating and maintaining independent evaluation space, while supporting ownership and mutual accountability, is particularly challenging for those managing evaluation teams in such settings.

<u>Meeting expectations:</u> Evaluation managers and teams highlighted that low capacities, the lack of qualified consultants (with both conflict and evaluation expertise) and the over-ambition of evaluation terms of references often made for unmanageable evaluation processes with extremely wide scopes. One evaluation (Sri Lanka) began with 92 evaluation questions! Expectations must be managed and agreement reached during the inception phase about what can (and cannot) be achieved through a single evaluation. In situations of "hot conflicts" (i.e. Afghanistan and Iraq) particularly expectations will have to be lowered, both because of the danger posed to evaluators and local partners and because of the highly political nature of working in such contexts.

Next steps
Policy lessons and key messages from this collection of evaluations are being compiled and disseminated widely as inputs to national dialogues and broader discussions on statebuilding, peacebuilding and aid effectiveness. It would be useful to compile the lessons from this series of evaluations to feed into policy discussions and inform development agencies work on peacebuilding, statebuilding and fragility. Likewise, it was suggested that a compilation of the methodological lessons and perhaps a complimentary case book of concrete examples of how the guidance is applied in different contexts, would also be useful for practitioners and evaluators.

Based on the outcomes of this event and other feedback received, an advisory team will oversee revision of the guidance by a consultant. The team will be made up of members and the secretariat of both networks and will work with two peer reviewers, one from INCAF and one from EVALNET. A final revised document will be shared with both networks at the end of June for consultation, with the aim of completing the approved final document in early autumn 2011. It will be presented to the EVALNET meeting 23-24 June 2011.