

## EVALUATING PEACE MEDIATION AND PEACEBUILDING in ONGOING CONFLICT

### Correct Use of the OECD Evaluation Criteria and Implications for Evaluations and Evaluators

June 2023

Peace mediation and peacebuilding are notoriously difficult endeavours. Many actors are definitely beyond the ‘control’ of the mediators or peacebuilders, very difficult to influence or even hard to reach. Most conflicts evolve into a set of interlocking conflict dynamics, with international, regional, national and local layers that are interconnected but also have somewhat different driving factors and key actors.

Yet donors, as well as mediators, want to assess, even ‘evaluate’ the ‘effectiveness’ and ‘impact’ of their actions. The question is: What is the reference, what are relevant criteria? Two main reference sources can be used: the OECD evaluation criteria, and what we have learned, from years of comparative experience, about practices that *increase the likelihood* of having some positive influence or impact.

This brief focuses on the OECD evaluation criteria, how the OECD advises they be used and how applicable they are to peace mediation and peacebuilding. It concludes with the observation that using these criteria to evaluate this type of actions, is *by no means only a matter of methodological competence*. That has implications for the choice of reviewers and evaluators.

## I. THE EVALUATION CRITERIA

### 1. The OECD Evaluation Criteria

The Development Assistance Committee (DAC) of the OECD first spelled out the key evaluation criteria (relevance, effectiveness, efficiency, impact, and sustainability) in 1991. It subsequently defined the terms in the 2002 *Glossary of Key Terms in Evaluation and Results Based Management*.<sup>1</sup> These criteria have since been the core reference for evaluating projects, programmes and policies in international development and humanitarian action. They also tend to be used for other types of interventions supported by international aid, such as governance reform and peace work.

In 2019, following an extensive review by the global evaluation community, the OECD DAC Network on Development Evaluation published revised definitions and principles for the use of now six criteria: relevance, coherence, effectiveness, efficiency, impact, and sustainability. In 2021, a further publication *‘Applying Evaluation Criteria Thoughtfully’* offered deeper reflections and considerations in their application. The need to provide further clarification and guidance on how to use them ‘thoughtfully’, suggests that this has not always been the case.

#### **Relevance**

Relevance as first criterion is not a coincidence. If an intervention is not that relevant to what matters most, then all the other performance criteria lose weight (see also Rogers 2012). In simple terms, relevance asks: is the intervention doing the right things?

---

<sup>1</sup> These are currently being revised with a new edition likely to be published soon.

### **Coherence**

Coherence as a second criterion is also meaningful because it directs us to a system perspective. Coherence asks: How well does the intervention fit? On the one hand, there is internal coherence: for an intervention with multiple components, are there internal linkages and cross-fertilizations that can make the whole more than the sum of its parts? On the other hand, and equally importantly, is the intervention complementary to, coordinated and/or harmonised with that of others, not only avoiding duplication of effort but again creating synergies that are more than the sum of the parts? In practice, this means that evaluators cannot just review the intervention they have been commissioned to evaluate but must assess interventions of others that pursue similar goals and objectives. Terms of reference and budgets should foresee this – which they seldom do.

### **Effectiveness**

Effectiveness criterion relates to the most closely attributable results. Here the evaluators explore whether the intervention is achieving, or expected to achieve, its objectives, particularly the more important ones, how, which factors were decisive, and whether there are equal or unequal results across groups? <sup>2</sup> The reference to ‘*most important results*’ has been introduced by the OECD as not all expected results are of equal importance; the overall appreciation/judgement of effectiveness must take this into account. Effectiveness is differentiated from ‘impact’, which examines higher-level effects and broader changes. (OECD DAC 2021:52, 64 ff).

### **Efficiency**

inquires whether the intervention delivers results in an economic and timely way. The guidance points out that timeframes may have to be adjusted to the evolving context, and that the cheapest intervention is not necessarily the best option. It does not put value-for-money as central as some donors tend to do. It also signals that efficiency, rather than being interpreted exclusively in cost-effectiveness or value-for-money terms, can be interpreted as *operational efficiency: how well has the intervention been managed?* This becomes particularly important in volatile environments, where appropriate adaptive management might be preferable over ‘sticking to the plan’.

### **Impact**

Impact inquiries into what difference, positive or negative, the intervention has made or is expected to make, *at the higher level*, beyond its more direct outcomes. Here we are looking at possible effects that are broader in scope, potentially transformative, at the levels of norms, institutions, behaviours of very influential actors, etc., and are therefore likely to be enduring. Impact assessment should not be done at the project level. (Jean & Ernstorfer 2014:3; OECD 2021:64 ff). Impacts may also not manifest themselves quickly; so when an evaluation takes place in the midst or at the end of an intervention (or its funding period) it will often only be able to speculate about eventual ‘impacts’.

### **Sustainability**

Sustainability asks whether positive results are likely to last? In other words: is the broader system, beyond the intervener, likely to be able to continue the positive results? Once more, this is a question that can only be answered with a longer time frame. The guidance is clear that this requires the consideration of relevant factors such as social factors and capabilities, political (and military) dynamics, institutional capabilities, finance etc. For an evaluation timed closely to the intervention, the answer will again be speculative, not firm: we are talking about the ‘*probability of continued long-term benefits*’. (OECD DAC 2019:12)

In practice then, reviews and evaluations will mostly concentrate on relevance, coherence, effectiveness, and efficiency. *Impact and sustainability are bigger and longer-term issues that are not directly linked to the quality of the intervention, and often beyond the practical scope of the evaluation exercise.*

The two most recent OECD guidance notes (2019; 2021) are telling in other ways: They deliberately talk about ‘intervention’, to avoid it being mentally framed as a ‘project’ with its specific connotations. It can refer to a programme but equally to an intervention that is largely designed as a process approach with allowance for emergence. ‘Emergence’ is not ‘emergency’: it means that not every step is predetermined but that a process-based intervention can respond to what comes up and follow on from there. The appreciation then does not rely on overly ‘*determined design*’, i.e. highly detailed pre-planned

---

<sup>2</sup> This has been introduced to encourage greater attention to equity among who benefits from the results.

interventions, but on the quality of ‘*process design*’, and of process management (the efficiency criterion).

## 2. The OECD Criteria in Relation to Peace Mediation and Peacebuilding

In 2012, the OECD/DAC guidance produced the first authoritative reference for evaluating peace activities and was explicit about its distinctive challenges, as the text box shows.<sup>3</sup>

### **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.

### **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.

### **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.

*OECD DAC 2012: 32-33*

Noteworthy is that the 2019 and 2021 publications occasionally mention conflict settings, but their reflections and guidance remain predominantly related to development and humanitarian actions. Contrary to the recognition, in 2012, of the higher degree of complexity facing peacework, certainly in ongoing open conflict, and the difficulty in ascribing causal links, this most recent guidance remains profoundly framed in terms of problems that can be addressed with known cause-effect relations (‘results-chain’; ‘causal pathway’), and where objectives of the intervention can be defined in a SMART manner (specific, measurable, attainable, relevant, timely). (OECD 2021:38). They do not really acknowledge that certain situations can be ‘complex’ (as understood in David Snowden’s Cynefin

<sup>3</sup> The difficulty of evaluation peace processes is echoed by experienced evaluators, e.g. “*In the peacebuilding field, evaluation is challenging because the work involves long time frames, complex fast changing dynamics and important human factors that appear intangible.*” (Woodrow et al 2017: 7)

framework), and that active conflict situations (as currently in e.g. Myanmar and Ukraine) can be very VUCA (volatile, unpredictable, complex and ambiguous) and this for a long period of time. In such contexts, it would be hard for any actor to have an intervention strategy that will ‘deliver’ smart results with a high degree of confidence, while questions of ‘impact’ and ‘sustainability’ would be entirely premature.

### **3. The Application of the Criteria to Peace Mediation and Peacebuilding in Complex Settings**

The following section discusses the applicability of the OECD/DAC evaluation criteria to peace mediation and peacebuilding in complex and volatile situations.

#### **a. Relevance**

The OECD/DAC guidance speaks about being relevant to the needs and priorities of intended beneficiaries, or the policies and priorities of institutions. For situations of protracted conflict, the relevance can better be determined with reference to a conflict- and peace analysis, at any level of a multi-layered conflict system.

#### ***Relevance and conflict- and peace analysis***

Good conflict- and peace analysis should try to differentiate between key drivers of conflict and of peace and contributing factors to each force. Contributing factors to conflict for example, could be the wide availability of arms, experiences of trauma in a population, a widespread lack of trust in the leadership, the lack of reliable information or the ability to differentiate more reliable information from rumour and deliberate misinformation etc.<sup>4</sup> These are important contributing factors: reducing them will improve the conditions for effective peace work on the drivers of conflict – but it does not by itself change these drivers. Good conflict- and peace analysis also includes ongoing analysis of the key actors and their supporters, and the power dynamics between them.

In practice, comprehensive and up-to-date conflict and peace analyses are not always available and when available not necessarily shared. They may also become impractical to produce: how to summarise 60 years of conflict in Burma/Myanmar, with deeper roots in fault lines created already during the colonial period? In the analyses that are produced, there can be significant weaknesses and gaps: for example, few will capture that there can be conflicting analyses, reflecting the different and contesting narratives of the main protagonists, which is precisely one contributing factor to ongoing antagonism (Van Brabant 2017).

A first observation must be that assessing the quality of a conflict- and peace analysis is not a methodological skill. It will require solid understanding of the context(s). Oftentimes, evaluators do not have this prior to starting the evaluation and need the time and commitment to seriously familiarise themselves with the context. Alternatively, they may have been working in that context and may then be relying on their own (possibly implicit) conflict- and peace analysis, the quality of which itself must be critically reviewed.

There are also choices to be made about what reference for ‘relevance’ is given most weight? Is it the policy and strategy of the donor in this conflict, or is it the evolving conflict dynamics? Sometimes, an intervention has been designed and is funded because of its fit with a donor policy, strategy, or priority. High relevance of the intervention in this regard is not necessarily reassuring: everything depends on what conflict and peace analysis those policies and strategies are derived from? Donor policies and strategies should not be left beyond the scope of the evaluation, as is often the case. But their relevance must be evaluated against credible conflict- and peace analyses. If intervention shows high relevance within the apparent conflict and peace dynamics but is not aligned with a donor policy or strategy, the implication is probably that the policy and strategy need to be reviewed, not the intervention.

---

<sup>4</sup> The key conclusion of a study of peacebuilding projects by four bilateral donors was a widespread ‘*strategic deficit*’. (Smith 2004). Many did things they claimed would contribute to peace but could not explain how their action had any influence on the drivers of conflict or at a broader level.

### ***Relevance and theory/ies of change***

Relevance can be further examined with reference to a ‘theory of change’. Theories of change are well known in the peace field (and other aid sectors), though often poorly practiced.

Firstly, they were originally developed following the frustration with the logical framework and its overly rigid causal pathways assumptions. It is not unusual however to now see them appear as ‘glorified logframes’ or constructed around a result-based framework, with the same overconfident cause-effect reasoning that is unfit for complex contexts. Logical and results frameworks can remain predicated on three deep seated assumptions that do not hold in open conflict situations: that the operating environment is fairly stable and therefore predictable; that the intervention is a factor with a fairly high degree of influence compared to other factors; and that the intervention managers have (and must have) a high degree of control over it. Theories of change can also be considered more for their usefulness for monitoring and evaluation, than from the perspective that for the peace actors is most difficult: planning in the face of high uncertainty.<sup>5</sup>

Secondly, theories of change only become interesting and potentially useful if they are highly contextualised. A generic theory of change is not good enough. The question should be about ‘*how do we think this state of affairs in this particular context can be changed to something more positive?*’ This question directs one once again to consider the underlying context and the conflict analyses and their quality. If the intervention works in different contexts, e.g. in different locations of a country, which each have to a degree a dynamic of their own, or if the intervention has different strands of work each focused on a certain factor of conflict- and peace, then to be really useful one may need theories of change for each. For example, how one could affect positive change in southwest and in northeast Syria may require very different tactics because the key actors and dynamics are different. The pathways towards change therefore may also differ. As another example, a programme component bringing together and fostering constructive dialogue between political parties and another programme component focusing on increasing women’s understanding of the peace process and federalism might both need their own theories of change to be tested in actual practice and evaluated separately.

Managers of complex peace programmes that operate at different levels and in different locations must work with multiple theories of change and adapt them based on experience of what happens. In practice articulating, in a proposal, separate theories of change for the different components/locations of an intervention, is generally too much detail to be realistic. Nor is evaluating each component or strand of a multi-facetted intervention against its underlying theory of change, usually feasible.<sup>6</sup> There is a risk then that the evaluator relies on her or his, possibly implicit, personal theory of change as a reference. That, however, should itself be the object of critical review.

### ***Relevance and theory of durable peace***

Less recognised than theories of change is the often-left implicit *theory of durable peace*. A theory of change focuses on how we think we can get from a negative to a constructive dynamic. A theory of peace asks what, in this country, would enable durable peace? Western actors oftentimes tend to differentiate between a ‘cold peace’ (low overt violence but probably much structural violence and continuing tension points beneath the surface) and a ‘just peace’ – based on rule of law that respects fundamental rights of all, addresses past wrongs and injustices etc., and where the general attitude to deal with ongoing or new disagreements and conflicts is negotiation, dialogue or acceptance of arbitration. Similarly, Western-funded peace efforts and initiatives tend to be inspired by the theory that a neoliberal model of economic development and political democracy are the best option for durable peace everywhere. That theory of durable peace is not necessarily shared by all conflict actors.<sup>7</sup>

### ***Relevance and adaptive management***

In volatile and complex environments, where there are so many actors and factors that influence the dynamics at different levels, it will be unclear at the outset which intervention (component) may turn

---

<sup>5</sup> See e.g. Corlazzoli & White 2017

<sup>6</sup> The problem is not methodological feasibility but the usual time limits most evaluations work under, as well as problems of memory and reliable recall by the intervener, if the theories of change were not documented.

<sup>7</sup> See e.g. <https://www.frontiermyanmar.net/en/why-the-wests-state-building-practices-in-myanmar-are-part-of-the-problem/>

out to be positively influential, or even whether the theory of change is appropriate. Even if it was, major changes in the operating environment, beyond the control of the intervention, may force it to change significantly. Keeping the intervention relevant may require adaptation, even significant deviation from the original plan. Therefore, simply continuing as per the originally designed and contracted plan could be a fundamental error that affects the relevance. Adaptation, even deviation from the original plan, can be an indicator of high-quality practice.<sup>8</sup>

### **b. Coherence**

When considering *internal coherence*, the first question for a peace intervention with multiple components is whether there are internal linkages and cross-fertilisations that can make the whole more than the sum of its parts. Appreciating this will require the evaluators to assess whether and how the components fit within the intervener's conflict and peace analysis, theory/ies of change and theory of durable peace. The second question is that of *external coherence*; is the intervention complementary to, coordinated and/or harmonised with that of others, and explicitly or implicitly creating larger synergies? In practice, this means that evaluators cannot just review the intervention they have been commissioned to evaluate but must assess interventions of others that pursue similar goals and objectives. Terms of reference and budgets almost never foresee this.

Donors who ask evaluators to review the complementarity between different interventions they co-fund, and provide the time to do so, may still miss the point. This is since *coherence under a donor policy or strategy* is secondary to *coherence of interventions in a specific location, around a certain theme, and/or with certain actors influential on the dynamics of conflict and peace*, irrespective of who funds them. For example, competing mediation efforts in the past have hampered progress in e.g. Darfur and Libya, and continue to do so today over Syria. Each competing mediation actor may be coherent in the different actions they fund, but the net result remains competing mediations that can be played against each other by conflict entrepreneurs, which undermines overall effectiveness.

However, while external coherence of an intervention can add to the collective effectiveness, evaluators need to recognise that peace actors may sometimes consciously decide to distance themselves from interventions led by others, because they consider them misplaced or likely to be counterproductive. Whether this is justified or not is a difficult assessment to make, and not one that can rely on methods only.

### **c. Effectiveness**

Mention has already been made of the failure of the latest OECD guidance (2021) to take note of its earlier (2012) acknowledgement that there are no known or predictable cause-effect relations in peace processes so that peace work is not linear. The OECD guidance also fails to distinguish clearly between *programme effectiveness* and *peace effectiveness*. (Jean & Ernstorfer 2014:3) Programme effectiveness asks whether the intervention largely took place according to the (contractually fixed) plan. In situations of complexity, the assumption that {more detailed design = greater effectiveness} is simply wrong. As understood here programme effectiveness is secondary to peace effectiveness, for which, certainly in contexts of high uncertainty, smart and timely adaptation may be the right actions, even if no longer in line with the original plan.

### **d. Efficiency**

The 2021 OECD/DAC guidance signals that efficiency does not necessarily refer to the most economic manner of achieving a result, but to the relevance and quality of the approaches followed to try and obtain peace effectiveness.

Understood in this manner, the efficiency criterion invites attention to the strategy and tactics of change agents – which is what peace actors seek to be – and the quality of the process management, not just the results. When it comes to influencing human behaviour, the ‘how’ you try to exercise or create ‘influence’ can be as important if not more decisive than ‘what’ you do. Many change processes fail, not because they did the wrong thing, but because they did it wrongly. It also allows a positive evaluation of repeated

---

<sup>8</sup> ‘Evaluators should consider relevance not only and the beginning and end of a programme, but how it has responded to changes over the course of its lifespan.’ (OECD 2021:40)

adaptations, to maintain relevance and the probability of effectiveness in a shifting context, even if such repeat adaptations add to the cost.

#### ***e. Impact and sustainability***

As understood *per the OECD usage of the terms*, attempting to evaluate impact and sustainability, certainly in ongoing active conflict and in relation to one specific programme, will remain a highly speculative exercise and is unlikely to be a good use of money. Many violent conflicts, like those in South Sudan, Nigeria, Mali, Yemen, Syria, Myanmar, Central African Republic, Democratic Republic of the Congo, continue for over a decade and more. Sustained peace remains elusive. Worse, even where there have been overarching peace agreements and large-scale violence has substantially diminished, years later their sustainability can appear shaky, as we observe for example in Bosnia-Herzegovina and Colombia. But that does not mean that all peace processes in these countries have been low on relevance and effectiveness, though it does raise questions that would be worth exploring with a very comprehensive review looking back over many years. That would certainly be a multi-donor and multi-team exercise. The questions of impact and sustainability also refer us to what has been called ‘perpetual peacebuilding’ (Paffenholz 2021). If ‘impact’ is used with a different understanding, which the OECD would call ‘effectiveness’, then the clear and common understanding of the term between the commissioners of the evaluation, the intervenors being evaluated and the evaluators, needs to be agreed upon.

## **II. IMPLICATIONS FOR EVALUATIONS AND FOR EVALUATORS**

### **1. Implications for Evaluations**

#### **Prioritise mid-term reviews over end evaluations and use the OECD/DAC criteria thoughtfully**

The long-term nature of peace processes, also after formal agreements are signed, and the almost certain necessity of reflective practice and adaptive management, mean that real-time reviews are more useful than end-of-project evaluations.

#### **Application of the criteria, challenges, and the scope of evaluations**

The 2019 and 2021 guidance are explicit that the six criteria should not be used mechanically but thoughtfully. There is no obligation for all evaluations to try and cover all of them, and other criteria can also be used. Criteria need to be chosen that are fit-for-context and fit the purpose of the evaluation. (OECD 2019:3).

The clarification of the criteria and how to use them, does not diminish the challenges for the evaluation of peace work, mentioned in the 2012 guidance. We know for example that mindsets can be influential in shaping people’s behaviours, so that sustained behavioural change requires changes in mindsets. Collecting ‘data’ about mindsets is not a straightforward exercise though.

The scope of the evaluation needs to be relevant for the peace objective and proportionate to the resources allocated to it. It is possible, for example, to assess an organisation’s peace intervention in a particular subregion of a country. But from a peace effectiveness approach, it might be more relevant to assess all (or at least the most important) peace interventions in that region, and how they connect to national level dynamics and top-down approaches to the same region. That, of course, would require consent of other interveners and their donors to be included in the evaluation. In the same vein, a donor can call for an evaluation of all its peace support in a particular context over e.g. five years. While that may be interesting for the donor and its accountability to its parliament and taxpayers, it does not define the scope on the basis of the conflict- and peace dynamics, which is ultimately where the question of effectiveness relates to. Evaluating the work of an organisation whose intervention consists of building and supporting collaborative networks between multiple actors which, between them, operate at different levels is feasible – but will require significant evaluation resources to cover the full spectrum of activities. Sometimes the evaluation of one or a few specific components of a multi-faceted programme can be more insightful.

## **Open inquiry evaluation**

Historically, the tendency has been to ask reviewers and evaluators to assess whether an intervention achieved its intended objectives or results. Framing an evaluation as a check whether the intervention has operated as per the determined design from possibly some years before, when the funding contract was established, misses the key point. A ‘yes’ answer may actually indicate that the intervener has been on automatic pilot and failed to adapt where needed, which may have reduced its relevance and effectiveness. (e.g. Peace Direct 2016:33-38). What it requests the evaluators to assess is whether a contract has been executed according to a past plan, not whether the intervention contributed to any meaningful change in the conflict- and peace dynamics. Evaluating against a preset design also focuses the attention (like predetermined indicators do) on certain aspects of a situation, with the risk that other relevant aspects or unintended effects (which may be positive or negative) remain out-of-field. Yet the DAC definitions of 2012 and 2019 are very clear that we are also looking for unplanned and unintended impacts.

A different evaluation approach, seen as particularly relevant for complex situations (e.g. Saferworld 2019, USAID 2021) is therefore to inquire in an open manner what changes actually took place and why, and only subsequently relate this to the intervention objectives. Methods like ‘Most Significant Change’ and ‘Outcome Harvesting’ are appropriate for this.<sup>9</sup>

## **Include donor practices in the scope of the review or evaluation**

How have the funders of the intervention affected it? Every practitioner knows how the priorities, terms and conditions, formats, procedures and requirements of donor administrations impact interventions, in an enabling or constraining manner. They are a potentially very influential factor on the relevance, efficiency and potential or actual effectiveness of the intervention. They cannot be left out of the picture.

## **2. Implications for Evaluators**

### **Evaluability**

Can the intervention be reviewed or evaluated in all its aspects? This is an important practical question to ask upfront. Many factors can limit the evaluability such as budget and time constraints; insufficient or inadequate documentation; lack of access to intervention activities (e.g. because of safety or security concerns or transport limitations) impeding the ability to observe; inability to speak with what would be key informants or with enough of them, or hesitations among those to speak freely; unwillingness of people to respond to surveys etc.)

Where evaluability is problematic, a modest but still useful form of critical reflection exercise is a Programme Quality Assessment. Unlike an evaluation, it does not try to assess results, impacts, contributions and attributions etc. Its focus is on examining the quality of the intervention logic and its management, or what has been called in this brief: how smart the navigation and navigators are. Areas to look into are:

- a. The quality of a conflict analysis contributing to program *relevance*;
- b. Clear and appropriate program goals;
- c. Programme strategy/ies and logic, including incorporation of structured reflective practices as part of the program design
- d. Well-formulated and plausible theory(ies) of change at different levels (macro – meso-micro) which are intentionally reviewed by the intervention managing team;
- e. Inclusion of conflict-sensitivity in design and implementation. (Ernstorfer et al 2016:19)

### **Reviewing or evaluating peace mediation and peacebuilding requires more than methodological skills and can never be a totally objective exercise**

While it is more straightforward to assess whether an intervention is relevant against the donor’s policy and strategy, assessing its relevance for the conflict-and peace dynamics will depend on the evaluators’

---

<sup>9</sup> There are also misunderstandings and misuses of Outcome Harvesting, e.g. when it is used to elicit the outcomes rather than the activities against a results-based framework. Correct use means starting from what happened in reality (or what people believe happened and why), not from pre-planned objectives or results. Guidance and examples can be found at <https://outcomeharvesting.net/> and <https://dgroups.org/groups/outcome-harvesting/outcomeharvesting/>



own (often implicit) conflict and peace analysis of that context, and their own (often implicit) theory of change – or how credible they find those of the peace actors whose intervention they are evaluating. The evaluation or appreciation of the tactical and possibly strategic choices peace actors have made is another matter of interpretation. With hindsight, it may be possible to conclude that certain choices were wrong as the desired outcomes did not materialise, but a fair appreciation acknowledges that peace actors do not act with the benefit of hindsight, but often in situations of high uncertainty.

The same goes for efficiency, if understood as the skillful management of a process in turbulent waters rather than a primarily cost-focused assessment. The ability to assess ‘coherence’ may not exist where the evaluators do not have the time or agreement to look at other peace efforts and how the one that are focused on related to those or not, and why. And while there is scope to assess ‘effectiveness’ in the sense of outcomes that can be plausibly linked to the intervention being evaluated, so many factors and actors influence what happens, that often the link will be one of contribution rather than fuller attribution. Even that may be difficult to properly appreciate, as giving the credit for a positive change to the actors one wanted to influence, may precisely have been one factor that led to the influence.

In short, evaluating conflict reduction, conflict transformation or peacebuilding in ongoing conflict situations with a high degree of uncertainty, requires much more than methodological skills. It certainly requires familiarity with the good practices derived from the collective learning, that will be summarized in another brief. Evaluators of peace mediation and peacebuilding must work themselves fairly deeply into the context. They need to appreciate that a peace intervention takes place in an environment full of antagonistic actors in a constant power dynamic with each other, who do not necessarily reveal their real interests and motives. And they need to be able to appreciate the tactical and strategic importance of how the peace actors position themselves in that arena, what roles they play and what level of profile they seek.

The last text box identifies several team characteristics that increase the probability of a review or evaluation team conducting a relevant and fair assessment.

#### **Essential Competencies and Characteristics of a Team Evaluating Peace Work**

- Willingness and ability to really work themselves into the context(s) with solid depth of understanding, including the somewhat ‘longer’ history which often plays an influencing role, if only in the narratives of protagonists
- Background in political or social sciences, with global and historical perspectives
- Familiarity with mediation and with peacebuilding learning, and preferably with actual peace practice
- Awareness of their own, possibly implicit, theory of durable peace, and theories of change which can introduce subjective bias in their appreciation of the intervention being reviewed/evaluated
- Understanding of the differences between working on complicated and on complex issues, and of appropriate review and evaluation references for each
- Awareness that their review/evaluation is an intervention, and needs to be done with great conflict-sensitivity
- Attention to how donor procedures and requirements can impact on the quality of intervention management
- The ability to speak normal everyday language to inquire into change processes, and then translate it back into the technical jargon that the international aid sector uses
- Competency in at least one of the important local languages
- Understanding of social interaction and communication norms in the context where the intervention (and the review/evaluation) takes place
- Self-awareness about gender dynamics and strong practical gender-sensitivity
- Practical faith-sensitivity
- Clear acknowledgment that evaluation involves a judgmental appreciation, which to a smaller or larger degree will also be subjective – and the responsibility that comes with such power

This is likely to require a team that is a mix of insiders and outsiders with gender balance.

Finally, reviews and evaluations are, themselves, an intervention. (Woodrow et al 2017:30) The questions that are asked and how they are asked, as well as the behaviours of the reviewers and evaluators, may affect how the interlocutors see the intervention and the peace actors– and potentially do harm.

## Selected Sources and References

No author 2016: Complexity-aware Monitoring. Programme cycle guidance. Policy, Planning and Learning – USAID

No author 2018: Complexity Theory. The Cynefin model. Clear Impact Consulting Group

Bryan, K. & P. Carter 2016: Contracts for Adaptive Programming. Overseas Development Institute

Byom, K., M. Ingram, A. Oakley & L. Serpe 2020: Adaptive Management. A practical guide to mitigating uncertainty and advancing evidence-based programming. PACT

CDA Inc. 2016: Designing Strategic Initiatives to Impact Conflict Systems. Systems approaches to peacebuilding. A resource manual

Church, Ch. & M. Rogers 2006: Designing for Results: Integrating monitoring and evaluation in conflict transformation programmes. Search for Common Ground

Corlazzoli & White 2017: *Change in Conflict, Security and Justice Programmes. Using theories of change for monitoring and evaluation.* Search for Common Ground

Cox, J. 2021: How does the Commissioning Process Hinder the Uptake of Complexity-Appropriate Evaluation? *Evaluation* 27(1) 32-56

Ernstorfer, A., I. Jean & P. Woodrow 2016: Thinking Evaluatively in Peacebuilding Design, Implementation and Monitoring. CDA Inc & Peacebuilding Evaluation Consortium

Honig, D. 2020: Actually Navigating by Judgment. Towards a new paradigm of donor accountability where the current system doesn't work. Center for Global Development, Policy paper 169

Jean, I. & A. Ernstorfer 2014: Alternative Approaches to Evaluating Peacebuilding Initiatives. CDA Inc  
Kleinfeld, R. 2015: Improving Development Aid Design and Evaluation. Carnegie Endowment for International Peace

Natsios, A. 2010: The Clash of the Counter-bureaucracy and development. Center for Global Development

OECD 2002: Glossary of Key Terms in Evaluation and Results-based Management

OECD DAC 2004: Lessons Learned on Donor Support to Decentralisation and Local Governance

OECD DAC 2012a: Evaluating Peacebuilding Activities in Settings of Conflict and Fragility. Improving learning for results

OECD 2012b: Improving International Support to Peace Processes. The missing piece

OECD 2019: Better Criteria for Better Evaluation. Revised evaluation criteria. Definitions and principles for use

OECD 2021: Applying Evaluation Criteria Thoughtfully

Paffenholz, T. 2021: Perpetual Peacebuilding: A new paradigm to move beyond the linearity of liberal peacebuilding. *Journal of Intervention and Statebuilding* 15(3):367-385

Peace Direct 2016: Putting the Local First. Learning to adapt when measuring change. London

Ramalingam, B., L. Wild & A. L. Buffardi 2019: Making Adaptive Rigour Work. Principles and practices for strengthening monitoring, evaluation and learning for adaptive management. Overseas Development Institute

Rogers, M. 2012: Evaluating Relevance in Peacebuilding Programmes – CDA Inc., Reflecting on Peace Practice Programme

Rogers, P. & A. Macfarlan 2020: Monitoring and Evaluation for Adaptive Management. Working paper series 1. Better Evaluation

Saferworld 2016: Doing Things Differently. Rethinking monitoring and evaluation to understand change

Saferworld 2019: Outcome Harvesting. Saferworld's approach

Sharp, S., C. Valters & B. Whitty 2019: How DFID can Better Manage Complexity in Development Programming. Overseas Development Institute

Smith, D. 2004: Towards a Strategic Framework for Peacebuilding. Getting their act together. PRIO

USAID 2017: Conducting an Evaluability Assessment for USAID Evaluations

USAID 2018: Program Cycle. Discussion Note: adaptive management

USAID 2021: Complexity Aware Monitoring. Discussion note (version 3)

Van Brabant, K. 2016b: Evaluative Thinking – beyond monitoring and evaluation. Global Mentoring Initiative

Van Brabant, K. 2017: Conflict Analysis. 13 weaknesses to avoid. Global Mentoring Initiative

Van Brabant, K. 2018a: Good Conflict-Sensitivity and Peacebuilding Donorship, Global Mentoring Initiative

Van Brabant, K. 2018b: Local Capacities for Peace: What questions to ask, Global Mentoring Initiative

Van Brabant, K. 2021: Has COVID Taught Aid Administrations about Adaptive Management? Global Mentoring Initiative

van Es, M. et al 2015: Theory of Change Thinking in Practice. HIVOS ToC Guidelines. HIVOS

Woodrow, P., N. Oatley & M. Garred 2017: Faith Matters. A guide for the design, monitoring and evaluation of inter-religious action for peacebuilding. CDA Collaborative Learning Projects & Alliance for Peacebuilding

*Global Mentoring Initiative is a values-based and purpose-driven consultancy and advisory group. Navigation360 is its subunit, specialised in conflict and peace dynamics.  
[www.gmentor.org](http://www.gmentor.org) & [www.navigation360.org](http://www.navigation360.org)*

**Creative Commons 4.0** You are free to share, copy and redistribute the material or to adapt— remix, transform, and build upon the material, under the following terms: You must give appropriate credit and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests GMI endorses you or your use.

**Cite as: Global Mentoring Initiative June 2023: *Evaluating the 'Effectiveness' and 'Impact' of peace mediation and peacebuilding. Correct use of the OECD evaluation criteria and implications for evaluations and evaluators. Begins***