

Stabilisation Issues Note Monitoring and Evaluation

Purpose

Issues Notes provide a quick introduction to stabilisation issues for UK Government practitioners working on stabilisation. They point the reader to more detailed information and guidance. This Issues Note gives readers a basic understanding of Monitoring and Evaluation (M&E), so that they are equipped to integrate M&E into planning and execution of a stabilisation intervention. It sets out issues and questions that the practitioner may face, and points to other, more detailed sources of information. This note should ideally be read in conjunction with other Stabilisation Issues Notes.

Key Messages

- Monitoring and Evaluation should be an integral part of stabilisation interventions to improve planning, maximise intended impacts and reduce negative unintended consequences.
- M&E in stabilisation environments poses particular challenges. This is because the stabilisation environment is complex and turbulent, has multiple actors operating concurrently, there are often political pressures to demonstrate progress, and because the security situation makes gathering reliable data very difficult.
- M&E should be based on an explicit theory of change that sets out the overall aims of the stabilisation intervention. The theory of change should explain how proposed activities will contribute to the desired outcomes and impact, the assumptions it is based on, major risks and possible unintended consequences.
- M&E efforts need to be embedded in stabilisation planning, management and delivery rather than conducted as a discrete or ad hoc activity. People responsible for M&E should form an integral part of the intervention/programme team, and sufficient resources must be allocated, though the framework must also be proportionate.
- Appropriate indicators (both quantitative and qualitative) must be identified at the planning stage to enable real time evaluation of outcomes and impact (or campaign effectiveness assessment, in military terminology).
- Regularly reviewing the validity of theories of change and contribution stories will help build understanding of changes in the stabilisation environment and enable adverse impacts (or second and third order effects) to be identified early.
- Imperfect data might still be 'good enough' for monitoring and evaluation of stabilisation interventions; expert judgment can play an important role in contextualising and interpreting data.

- To avoid the risk of bias it is desirable to use information from a wide range of sources; it will however remain important to prioritise data collection in order to prevent information-overload.

What is Monitoring and Evaluation?

1. Monitoring and evaluation is an essential part of the planning cycle for any project, programme or intervention, enabling learning and adaptation, as well as accountability and broader lesson learning. It is defined here following the OECD DAC:

- Monitoring refers to systematic collection of data and information to provide those involved in an intervention or programme with adequate information to monitor progress against objectives and outcomes. It is different from reporting in that monitoring refers to the collection and interpretation of information and not the delivery of information.
- Review and evaluation are similar. An evaluation is a systematic and objective assessment of a completed intervention. Evaluation determines the relevance, appropriateness, and fulfilment of objectives. It looks at efficiency, effectiveness, impact and sustainability. An evaluation should enable decision makers to judge the relative worth of a programme and draw lessons for other and future programmes.
- A review tends to be less in-depth than an evaluation and can (and should) be done more regularly and throughout an intervention.

Why is Monitoring and Evaluation of Stabilisation important?

2. Monitoring and evaluating is particularly important in stabilisation environments because they are often complex, unpredictable and characterised by a lack of information which can lead to plans quickly going off track.

3. Firstly, as a planning and management tool it helps planners to establish common understanding of goals and of how inputs and activities are expected to lead to the realisation of those goals and what risks exist. Used regularly throughout the life of an intervention, it allows the outcome of activities to be tracked, and an assessment to be made of whether they are contributing to the strategic objectives of the intervention.

4. It also helps to identify and correct for the inevitable gaps and flaws in initial stabilisation plans, increasing understanding as the intervention unfolds. This is particularly important in stabilisation environments because of the complexity and volatility of the environment, and is compounded by speed of response commonly demanded as well as the number of different actors frequently present, each with a different understanding of the problem and how it should be solved. Regular reviews of progress then guide programming, both to ensure that plans remain on track and resources are being allocated in the most effective way (or to adapt and reallocate if not), and to reduce the likelihood of inadvertently doing harm.

5. Secondly, monitoring and evaluating is important in order to ensure accountability, gathering information systematically and asking the right questions about what is being done and why, and whether it is working. However, the need for upwards accountability must be balanced with the importance of an open and honest management culture in which mistakes can be quickly identified and corrected, without (necessarily) blame. Equally, care should be taken not to distort M&E systems as a result of pressure to generate 'good news' stories; communications functions should be separate from M&E.

6. Finally, monitoring and evaluation helps to generate evidence-based lessons which may be applicable to other programmes/interventions over the longer-term.

7. A good M&E system should be able to produce information that can be used for all three, but the purposes are distinct and it may be necessary to create some form of separation in order to ensure that a safe space is created to allow the first of these – enabling real-time learning and adaptation – to be effective.

What are the main challenges?

8. Stabilisation environments present particular challenges that make monitoring and evaluation especially difficult:

- Stabilisation environments tend to be complex and turbulent, progress is generally non-linear and the impact of activities hard to predict. This makes it difficult to establish reliable baselines and to distinguish between strategic shifts and shorter-term fluctuations, as well as to attribute cause and effect.
- The rapidly evolving and sometimes unpredictable nature of these environments, as well as the tendency to have to act quickly, means that planning – including the design of an M&E framework – is often done based on limited information and faulty assumptions.
- There are often multiple actors (local, national and international) undertaking a range of concurrent activities with different underlying logics over different time horizons.
- Much of the most relevant information to measure stabilisation progress – for example perceptions, relationships and behaviours – is qualitative in nature, can be challenging to measure and requires social/political analysis.
- Interventions are often politicised, which can create perceived imperatives to demonstrate tangible outputs and progress quickly, at the expense of longer-term impacts.
- Insecurity can make it difficult to ask the right questions to the right people, and to interpret the answers, or even place those collecting data and/or responding to questions in danger.

Who should take part in Monitoring and Evaluation of Stabilisation?

9. Involving a range of stakeholders in monitoring and evaluation can help to build common understanding of the objectives and progress of an intervention, as well as of the environment. Those responsible for implementing programmes should ideally be involved in the design of the monitoring and evaluation framework, and take part in activities such as regular reviews in order to ensure that the framework is relevant and useful, and that its findings are acted upon. Senior leaders/managers should show clear interest in and commitment to monitoring and evaluation, and must take ownership of the results – even when they are negative.

10. However, specific roles and responsibilities should be clarified. Specific technical skills will be needed in particular tasks such as methodology design and data analysis, and – particularly in insecure environments – it may be necessary (and more appropriate) to rely upon others to gather data. Those carrying out reviews and evaluations should also maintain some degree of independence from those responsible for results.

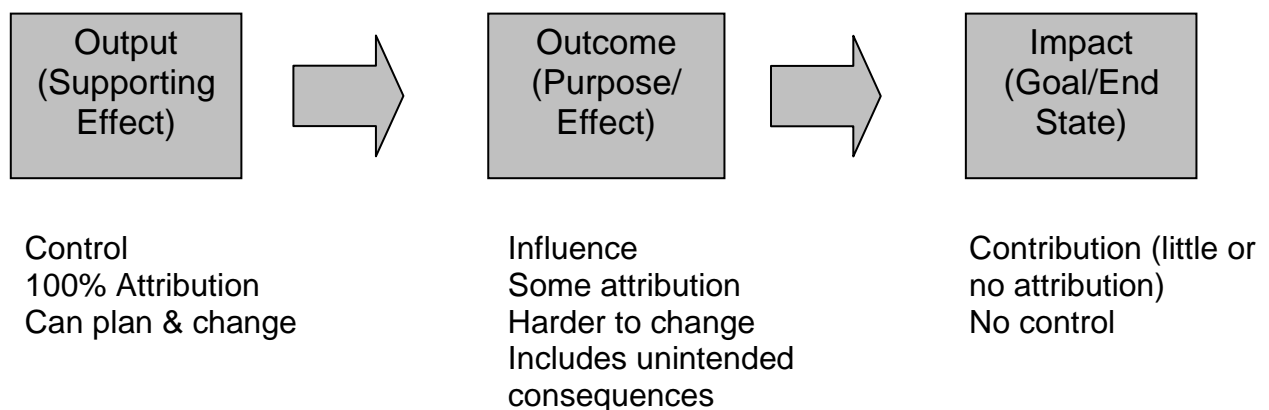
11. In addition, independent evaluations and audits are also likely to be needed and must in most cases be carried out by external teams.

How should Stabilisation be Monitored and Evaluated?

12. Monitoring and evaluation needs to be fully embedded into any stabilisation programme or intervention, built in at the outset and properly resourced throughout.

Articulate the theory of change

13. Any planning process should be based upon a ‘theory of change’ that sets out why certain activities are expected to lead or contribute to a particular desired outcome. Designing a monitoring and evaluation framework at the outset helps to clarify this thinking, and to make it more tangible. Common means of capturing the results of this process is in a results chain / logical framework, covering impact/goal, outcome/purpose, output and input, and assumptions; a risk matrix is also produced, normally separately but linked. Military planning process produce a similar result – with an end state, effects and supporting effects – but do not tend to capture the supporting logic to the same extent. In both cases, the level of control exercised by those implementing the intervention decreases in the move from inputs and outputs through to impact.



14. Theories of change in stabilisation will inevitably be far more complicated than a simple linear output-outcome-impact relationship, and will include areas of uncertainty and risk. Ideally, the articulation of the theory of change should include the following main areas – which a longer narrative produced in addition to the logframe or military schematic can help to flesh out :

- The objectives (in terms of desired impact/end state/goal and outcome(s)/effect(s)/purpose) of a programme/intervention
- The inputs/activities and outputs expected to realise those objectives
- Key assumptions underpinning the programme/intervention (highlighting critical causal links and areas of particular uncertainty)
- Relationships/interdependencies with other related programmes/actors
- Major risks (potentially undermining the success of the intervention)
- Other potential unintended consequences/harms, which can include
 - Ending up in the wrong place (not achieving the goal/end state due to faulty assumptions/weaknesses in the plan, or to unanticipated events)
 - Unexpected outcomes, both harms and benefits (directly due to the intervention) – some of which may not affect whether or not the desired impact/goal/end state of the intervention is achieved but nonetheless matter in and of themselves
 - Unintended impacts on other programmes and priorities, locally and/or at the strategic level (as secondary impacts of the intervention)

15. Moreover, there may not be one single theory of change – indeed it is likely that the reality will be more complex, particularly where different actors (civilian and military, local and international, non-state, inter-agency, multinational etc) are involved and acting independently. But the process of jointly planning and articulating theories of change should at least help to establish a deeper common understanding of objectives, to make visible implicit assumptions and beliefs about why change occurs, and reduce the likelihood of one intervention undermining the other.

Identify Indicators (Measures)

16. Good indicators should capture the ‘why’ as well as the ‘what’. This means they need to go beyond simple output measures (e.g. numbers of forces trained and equipped), and also gather information about quality and about the higher-level consequences of the activity (e.g. whether human rights abuses are decreasing as a result of training, and as a consequence legitimacy of the Government is increasing). They are needed at impact/goal, outcome/purpose and output levels – to allow attribution or at least contribution to be identified. In military planning these are called measures of success, effectiveness and performance, but are otherwise similar.

17. Stabilisation indicators (or measures) will need to address changes in perceptions, attitudes and relationships. These are, by their nature, subjective and difficult to measure, and will need to draw upon social and political analysis. In addition, it may be helpful to try to corroborate them by ‘hard data’ or measures of action to round the picture (for instance

balancing perceptions of confidence in the police with data on reported incidents and convictions). They should address both intended and unintended consequences. The latter can be addressed in part by monitoring risks identified during the planning process; but it may also be helpful to monitor potential negative outcomes even if they do not undermine the objectives of the intervention but where they are nonetheless significant.

18. Indicators may also need to address different timescales, including both near-term/intermediate and longer-term goals as progress may be slow and otherwise imperceptible. While they should be tailored to the specific intervention, and some may need to change as the plan does, at least a core set should remain the same – even if not perfect – to enable trends to be tracked over time. They may also need to be disaggregated to identify differences between demographic groups and/or between regions, particularly where inequalities are a driver of instability. Country-level indicators such as corruption or governance indices may therefore be useful, but will not be sufficient.

19. ‘Early warning’ thresholds (e.g. numbers of or levels of violence in protests) may be helpful, to allow problems to be identified and responded to early, particularly in volatile environments. Triggers may be set to reflect changes in trends or strange patterns of activity rather than absolute numbers.

20. It is important to keep in mind who will use the M&E framework and what for, and therefore what evidence will be ‘good enough’. For instance, if M&E is expected to support decisions over allocation of resources (money, deployment of troops etc) or where the consequences of getting it wrong are severe, a higher standard of evidence may be needed than if the indicators were to assess progress of a small and low-risk project.

21. It is also important to consider how realistic the indicators are – whether data is available or can be safely gathered at reasonable (proportionate) cost to support them, whether it can be routinely collected over time, and how reliable it is. Evidence needs to be actionable and sufficiently robust to support decision-making, and the number of indicators kept to a manageable level. Proxy measures such as availability and cost of ‘luxury goods’ in the market or personal investments such as installing TV aerials may be useful, depending on the standard of evidence that is needed.

22. Some good sources of possible indicators are listed in the Annex, and include DFID’s *Interim Guidance on Measuring and Managing for Results in Conflict-Affected and Fragile States*, *Measuring Progress in Conflict Environments* (MPICE), and *Development of a Basket of Conflict, Security and Justice Indicators*.

Gathering Data

23. Indicators need to be supported by reliable data. Common methodologies for gathering data relevant to stabilisation include through surveys, focus groups and interviews, as well as observation. These can be difficult in stabilisation environments because of the security situation. Local organisations can sometimes be used – although care must still be taken not to put either researchers or those interviewed in danger. Any information provided must be handled with absolute confidentiality.

24. In addition, insecurity and political sensitivities – as well as, in some cases, expectations of material gain – will affect responses to questions. Who is asking the question (a familiar face or a foreigner with a gun) will inevitably change the answer that is given. Even with the best methodology, some bias is unavoidable, and needs to be understood and factored into the interpretation of results so far as possible, with different sources and types of data ('hard' and 'soft') used to corroborate results.

25. Secondary sources – including information produced by NGOs, international partners, academics/research institutions – can also be used. This helps to fill gaps where information cannot be gathered directly (and is more efficient than duplicating efforts), and is useful in order to triangulate information gathered directly.

Analysing and Interpreting Data

26. Analysis to support reviews and evaluations should focus not only on whether desired outcomes have been achieved, but also ask why and what other unintended consequences have occurred, and changes in the environment have occurred that might affect the plan. Data should be used to review contribution stories/theories of change and assess whether or not they remain valid. Questions it may be helpful to ask include:

- Have activities been carried out properly and outputs delivered?
- Have outputs led to outcomes being achieved and has the impact been as expected?
- If not why not: e.g. were assumptions in the plan wrong, have any of the risks materialised, have other unexpected events taken place or has the environment changed?
- Have there been other unintended consequences as a result of the intervention?
 - has it done harm
 - have other priorities/objectives been undermined
 - have there been positive unintended consequences

27. Measuring trends over time and comparing across different areas is likely to be more useful than looking at specific numbers. Baselines will have been hard to establish, and may not be reliable; they should not be seen as absolutes but as comparators for trends. Particular findings may simply be an indication that more research is needed rather than giving an explanation themselves – showing that something is going on, even if it is not immediately obvious what.

28. Data needs to be analysed with the benefit of expert opinion and judgement to ensure it does not stand alone but 'tells a story'. Local knowledge and understanding is essential, and engaging those involved in programme implementation will help to explain and contextualise findings. For instance, perceptions of security measured on a particular day may have been affected by a specific, one-off incident rather than representing a general trend. 'Expert panels' are another possible means of verification and contextualisation.

29. Results can be presented in different ways. Popular methods include dashboards, scorecards and traffic lights to show progress against a handful of indicators. These can

be unhelpfully reductionist and should always be accompanied by a narrative explaining why certain trends or results have been seen, and additional information to highlight any unintended outcomes, positive or negative.

Reviewing and Adapting

30. In addition to continuous or regular monitoring activity, more substantive reviews should take place regularly throughout the life of the intervention, using the data and analysis produced by the M&E system to enable discussion and debate, and to support decisions on adapting the strategy, plans and allocation of resources. These should be led and owned by senior leaders/managers, with the participation of key stakeholders and those responsible for implementation. Organisational culture is critical to enable open and honest discussions.

31. In early stages of a stabilisation intervention, as more is being learnt about the environment and about the impact of activities, it is likely that plans (along with the theory of change) will alter more frequently. This may mean introducing new indicators, while retaining at least a core set of consistent ones to ensure that change over longer periods can be tracked, even if they are not perfect; the temptation to change indicators on a regular basis must be resisted. It is, though, entirely legitimate for a review process to identify areas of uncertainty in a plan/theory of change, and ahead of the next review point.

32. By later stages in the intervention, uncertainties within the plan should have been reduced and the focus be more on maintaining its course and identifying significant changes in the external environment.

Evaluating & learning lessons for other interventions

33. In-depth evaluations may be less useful during the life of the intervention, due to the time delay, but will still be needed for accountability reasons and are useful in learning broader lessons for the future. They will in most cases need to be carried out by people separate from the implementation of the intervention. However, they should nonetheless still be able to draw upon much of the same data used for M&E during the intervention; if this is anticipated in advance it may have implications for the quality and quantity of evidence gathered.

Resources and further information

- UK Approach to Stabilisation
<http://www.stabilisationunit.gov.uk/>
- DFID Interim Guidance: Measuring and managing for results in fragile and conflict-affected states and situations
- Guidance on Evaluating Conflict Prevention and Peacebuilding Activities, OECD, 2008, http://www.oecd.org/secure/pdfDocument/0,2834,en_21571361_34047972_39774574_1_1_1,00.pdf
- Designing for Results: Integrating Monitoring and Evaluation in Conflict Transformation Programmes, Search for Common Ground, 2006, http://www.sfcg.org/programmes/ilr/ilt_manualpage.html
- Monitoring Peace Consolidation: United Nations Practitioners' Guide to Benchmarking, UN, 2010, www.un.org/peace/peacebuilding/pdf/Monitoring_Peace_Consolidation.pdf
- Measuring Progress in Conflict Environments (MPICE): Metrics Framework, USIP, <http://www.usip.org/publications/measuring-progress-in-conflict-environments-mpice-0>
- Code of Best Practice for the Use of Measures of Effectiveness (MoE) to Support Campaign Assessment, DSTL, 2008 [available from DSTL on request]