# 3

# From M&E to monitoring and learning

"Would you tell me, please, which way I ought to go from here?" said Alice.

"That depends a good deal on where you want to get to," said the Cat.

"I don't much care where–"

"Then it doesn't matter which way you go"

*Alice in Wonderland* by Lewis Carroll

Lewis Carroll points to an important consideration when embarking on a learning exercise: if we want to prove or improve our work then we need to be able to describe clear intentions to direct our learning. We need to care where we are going. Without clear intentions, we are at liberty to define success in any way we like. This may sound appealing to some but is more likely to result in repetitive circles than learning.

The impact of academic research is traditionally evaluated via peer review to assess quality, relevance and accuracy; and citation analysis to assess uptake and reach. While both of these are important, neither helps us discover what influence the research may have had on policy (assuming it had an intention to do so), whether the research was worth undertaking and hence how to make it more effective. All we learn is how to make our research more attractive to other researchers!

Traditional M&E approaches – which rely on a simple feedback model with predefined indicators, collecting data and assessing progress towards pre-set objectives – are simply not adequate in the context of policy-influencing interventions. As explored in Chapter 1, many of the results we are looking for cannot be projected ahead of time in a linear fashion. The reality of the distributed capacities, divergent goals and uncertain change pathways that pervade many policy contexts means measuring progress along a predefined course is insufficient for monitoring.

Effective M&E requires a careful combination of sensing shifts in the wider context (policy, politics, economics, environmental, social), monitoring relationships and behaviours of diverse actors, weighing up different sources of evidence, being open to unexpected effects and making sense of data in collaborative enquiry. This kind of monitoring may seem challenging but it doesn't have to be. ROMA aims to shift the emphasis from evaluation and more to 'sense-making' of monitoring information. This fits into current management practices to ensure decisions on responding to an unpredictable situation are evidence-based and widely owned.

The purpose of this chapter is to provide a practical monitoring approach that builds reflective and evaluative practice into the work of influencing policy, to support decision-making and demonstrate progress. The chapter is split into three: the first part describes why and what to monitor; the second introduces practical options for embedding and carrying out this kind of approach; and the third studies how to make sense of learning and decision-making. The principles that underlie this approach can be summarised as:

**Appropriate to purpose, scale and context.** In ROMA, the primary driver for monitoring is the users and how they will use the data and insights. But scale and context are also determinants. A small-scale intervention will require much lighter monitoring than a long-term, multi-strategy intervention. (If you are not sure of the scale of the intervention at this stage, Chapter 2 will guide you through the planning process.) As with context, simple problems will require only routine monitoring and performance management, whereas problems exhibiting one or more signs of complexity will need more sophisticated, responsive and multi-purpose monitoring systems. (If you are unsure about the level of complexity, Chapter 1 will introduce you to three clear signs to look for.)

**Defines realistic results within the sphere of influence.** The influence of an intervention has a definite limit based on resources, time, reach, politics etc. Beyond the sphere of influence is the sphere of concern, which is where the results that really matter lie (such as better education, quality health care, secure livelihoods). However, you have to rely on others to influence these results. ROMA considers only results within your sphere of influence. These are the ones that can be measured and can guide strategy and engagement. The planning stages in Chapter 2 as well as the monitoring areas and measures in this chapter are used to define the intervention and its sphere of influence. They point to the priority areas to monitor.

**Focuses on actors and graduated change.** Much policy-influencing work revolves around people. It follows that monitoring policy influence should also revolve around people. In ROMA, an intervention is monitored through its effect on key stakeholders – those people or organisations within the sphere of influence of the intervention and whom the intervention seeks to influence directly or indirectly. ROMA recognises that effects can come in many guises and it is important to be able to pick up a broad spectrum – the simple, immediate responses that show you are on the right track as well as substantive commitments that indicate you are close to your goal.[10]

**Reasoned judgement about statistical significance.** ROMA is an inductive approach that seeks to generate evidence that can be used to increase our understanding of our effect on policy. It does not seek to determine a statistical, numerical measure of policy influence.

---

10. This is an adaptation of the concepts of 'boundary partners' and 'progress markers' in OM.

**Embeds learning in practice.** ROMA has two key strategies to keep it focused on learning: it builds on systems and practices that already exist; and it balances ongoing data collection with discrete studies. This data collection, with specific enquiries to cover the depth or breadth required, can be carried out by the intervention team but also commissioned from specialist researchers or evaluators.

**Building on collaboration and engagement.** Policy-influencing is an inherently collaborative exercise. It often requires bringing together a range of expertise and perspectives on a problem, building close relationships with influential people. ROMA is especially useful in this context as it helps teams learn together through shared monitoring priorities and opportunities for learning. The greater the engagement with key stakeholders, the greater the opportunity to collect and use meaningful data. However, ROMA is just a tool and doesn't achieve anything alone: that depends on the effort and commitment that users put into it.

# 1. WHAT TO MONITOR AND WHY

## Monitoring for learning and accountability

Broadly, the purposes behind M&E are usually viewed in terms of learning (to improve what we are doing) and accountability (to prove to different stakeholders that what we are doing is valuable). But we need to be more specific. Below is a list of nine purposes[11] that summarises different motivations and uses of M&E. Each will involve different elements of learning and accountability in a way that recognises the importance and interconnectedness of both rather than setting them in competition with each other. The first five purposes pertain to managing the intervention; the last four could be part of the intervention itself as strategies that directly contribute to the overall goal.

### List of nine learning purposes

**1** **Being financially accountable:** proving the implementation of agreed plans and production of outputs within pre-set tolerance limits (e.g. recording which influencing activities/outputs have been funded with what effect);

**2** **Improving operations:** adjusting activities and outputs to achieve more and make better use of resources (e.g. asking for feedback from audiences/targets/partners/experts);

**3** **Readjusting strategy:** questioning assumptions and theories of change (e.g. tracking effects of workshops to test effectiveness for influencing change of behaviour);

**4** **Strengthening capacity:** improving performance of individuals and organisations (e.g. peer review of team members to assess whether there is a sufficient mix of skills);

**5** **Understanding the context:** sensing changes in policy, politics, environment, economics, technology and society related to implementation (e.g. gauging policy-maker interest in an issue or ability to act on evidence);

**6** **Deepening understanding (research):** increasing knowledge on any innovative, experimental or uncertain topics pertaining to the intervention, the audience, the policy areas etc. (e.g. testing a new format for policy briefs to see if they improve ability to challenge beliefs of readers);

**7** **Building and sustaining trust:** sharing information for increased transparency and participation (e.g. sharing data as a way of building a coalition and involving others);

**8** **Lobbying and advocacy:** using programme results to influence the broader system (e.g. challenging narrow definitions of credible evidence);

**9** **Sensitising for action:** building a critical mass of support for a concern/experience (e.g. sharing results to enable the people who are affected to take action for change).

---

11. The nine purposes originate from Irene Guijt's work (see Guijt, 2008).

There are two very practical reasons for considering these learning and accountability purposes. First, making the purposes explicit directly links monitoring to the programme objectives and makes it clear to everyone involved why monitoring is important. This may be about informing stakeholders what is being done and what the effect is in order to sustain support for the intervention, or it could be to improve the ability of the team to effect the desired change. If this link is not clear then motivation for monitoring will likely decrease and it will be difficult to maintain participation and quality.

Second, each purpose will have different information requirements, different times and frequencies at which information is needed, different levels of analysis, different spaces where analysis takes place and information is communicated and different people involved in using the information. Clarifying the learning and accountability priorities can help thread these elements together to form a monitoring system embedded in existing organisational practices.

Table 8 presents a set of questions that can help decide the priority learning and accountability purposes. It also suggests possible users of the monitoring information gathered for each purpose. This helps when thinking through who needs to be engaged and what specifically they will need. Table 9 is a tool you can use for planning your monitoring by indicating the priority purpose(s) and describing where and when information is needed and who needs to be involved. The next step is to decide what information is required; this is described in the next section.

## Table 8: Prioritising users and uses for monitoring information

| Purpose | Key questions | Example users |
|---|---|---|
| Being financially accountable | Is money being spent where it was agreed it would be spent? <br> How does the funder define value for money? | People involved in managing, governance, funding |
| Improving operations | Are activities being implemented according to plan? <br> Is there a need for improving or redesigning activities? <br> Is immediate feedback available? | People involved in managing, implementing |
| Readjusting strategy | Are strategies leading to expected short-term changes? <br> Are short-term changes leading to expected longer-term changes? <br> Is the selection of key stakeholders still relevant and viable? <br> Are the objectives still appropriate? | People involved in managing, governance |
| Strengthening capacity | Is the team working effectively? <br> Does the team need new skills for effective implementation? <br> Are new systems or processes required? | People involved in managing, supporting implementation |
| Understanding the context | Is the intervention operating in a particularly unstable context? <br> How is the external political, economic or organisational context changing? <br> To what extent will these changes affect the intervention? | People involved in managing, in implementing; partners and stakeholders |
| Deepening understanding (research) | Do you have sufficient evidence to back up influencing activities? <br> Are there parts of your theory of change you are unsure about or don't have enough knowledge on? <br> Are you experimenting with innovative interventions? | People involved in managing, in implementing |
| Building and sustaining trust | Are there strategic partnerships, networks or coalitions that need building? <br> Is there information that can be shared that will help this? | Partners, allies and stakeholders |
| Lobbying and advocacy | Is there a need to influence policy outside the core policy objective – e.g. on research process, forms of evidence, viable interventions? | General public, specific audiences, partners and allies |
| Sensitising for action | Is there a need to build critical mass around this issue and enable others to support you in influencing? | General public, specific audiences |

## Table 9: Template for prioritising learning and accountability purposes

Short description of the policy-influencing goal and outline strategy

| Purpose | Priority | What information is needed? | Who will use information? | When and where is information needed? |
|---|---|---|---|---|
| Being financially accountable | | | | |
| Improving operations | | | | |
| Readjusting strategy | | | | |
| Strengthening capacity | | | | |
| Understanding the context | | | | |
| Deepening understanding (research) | | | | |
| Building and sustaining trust | | | | |
| Lobbying and advocacy | | | | |
| Sensitising for action | | | | |

Each of the purposes listed above will require different types of information about the intervention and its environment. For example, 'financial accountability' will require accurate information about the quality and quantity of what has been done and the resources used; 'understanding the context' will entail knowing about the people in charge of the policy area and their incentives; and 'strengthening capacity' will need information about performance of team members and partners, and the competencies required for the intervention.

# Monitoring for strategy and management

As well as monitoring for learning and accountability, monitoring also helps you ensure you remain headed in the right direction. There are six levels at which you can monitor:

- Strategy and direction (are you doing the right thing?)

- Management and governance (are you implementing the plan accurately and efficiently?)

- Outputs (do the outputs meet required standards and appropriateness for the audience?)

- Uptake (are people aware of, accessing and sharing your work?)

- Outcomes and impact (what kind of effect or change did the work contribute to?)

- Context (how does the changing political, economic and organisational climate affect your plans?)

For each of these levels, there are different measures you could consider monitoring. The full range of measures is presented below: for each of the levels, the measures are presented in a rough order of priority. You will probably already be working with many of them, but it is useful to go through the list below to see whether there are others you may need to consider.

## Strategy and direction

For many people working to influence policy, the choice of interventions will depend on the theory of change. Many start by making their theory of change explicit. This not only helps ensure a sound strategy in the first place, but also enables regular review and refinement at a strategic level. Practically, it helps identify key areas for monitoring and baseline data collection. Regardless of how a theory of change is presented, it can be assessed by questioning the following features:

**1** How the theory describes the long-term change that is the overall goal of the intervention: is the desired long-term change still relevant?

**2** How the theory addresses context: Is the strategy still appropriate for the context? Has the context changed significantly: does the strategy need to change?

**3** The assumptions about how change may occur at any point in the theory, and about the external factors that may affect whether the interventions have the desired effects: are the assumptions about policy change holding true? Has anything unexpected happened?

**4** How the theory assesses the different mechanisms that could affect the long-term change: is the assessment of the mechanisms affecting policy change still valid?

**5** What interventions are being used to bring about long-term change? Is there the right overall mix of interventions? Are the interventions having the desired effect, demonstrating movement in the right direction?

## Management

Management monitoring can be simplified down to recording what is being done, by whom, with whom, when and where. A systematic record of engagement activities can help make sense of the pathways of change later on.

Management monitoring can also involve assessing whether the most appropriate systems are in place, the best mix of people with the right set of skills are involved and the intervention is structured in the most effective way. This is particularly important when strategic policy-influencing introduces new ways of working for an organisation. Included in this is the regular assessment of the monitoring and decision-making processes themselves.

**1** Management and governance processes: how do organisational incentives help/hinder policy-influencing? Is learning from the team leading to improved interventions? Is the team working in a coordinated, joined-up way?

**2** Implemented activities: what has been done? When and where was it done? Who was involved?

**3** The mix of skills within the team: given the strategy, what capacity/expertise needs to be developed or bought in?

**4** Capacity or performance of individual team members: how are team members, contractors and partners performing at given tasks? What difference has training/capacity-building made?

## Outputs

Outputs are the products of the influencing intervention and communication activities. Policy briefs, blogs, Twitter, events, media, breakfast meetings, networks, mailing lists, conferences and workshops are all potential outputs.

It is not enough to just count outputs: quality, relevance, credibility and accessibility are all key criteria that need to be considered.

**1** Quality: are the project's outputs of the highest possible quality, based on the best available knowledge?

**2** Relevance: are the outputs presented so they are well situated in the current context? Do they show they understand what the real issue is that policy-makers face? Is the appropriate language used?

**3** Credibility: are the sources trusted? Were appropriate methods used? Has the internal/external validity been discussed?

**4** Accessibility: are they designed and structured in a way that enhances the main messages and makes them easier to digest? Can target audiences access the outputs easily and engage with them? To whom have outputs been sent, when and through which channels?

**5** Quantity: how many different kinds of outputs have been produced?

## Uptake

Uptake is what happens after delivering outputs or making them available. How are outputs picked up and used? How do target groups respond? The search for where your work is mentioned must include more than academic journals – for example newspapers, broadcast media, training manuals, international standards and operational guidelines, government policy and programme documents, websites, blogs and social media.

Other aspects to consider include the amount of attention given to messages; the size and prominence of the relevant article (or channel and time of day of broadcast); the tone used; and the likely audience.

Secondary distribution of outputs is also important. The most effective channels may be influential individuals who are recommending the work to colleagues or repeating messages through other channels: it is important to capture who they are engaging with and what they are saying. Finally, direct feedback and testimonials from the uses of your work should be considered.

The following are the results areas for uptake:

**1** **Reaction of influential people and target audiences:** what kind of feedback and testimonials are you hearing from influential people? How are they responding to your work?

**2** **Primary reach:** who is attending events, subscribing to newsletters, requesting advice or information?

**3** **Secondary reach:** who are the primary audiences sharing your work with and how? What are they saying about it?

**4** **Media coverage:** when? Which publication(s)/ channel(s)/programmes(s)? How many column inches/minutes of coverage? Was it positive coverage? Who is the likely audience and how large is it?

**5** **Citations and mentions:** who is mentioning you and how? For what purpose: academic, policy or practice?

**6** **Website/social media interactions:** who is interacting with you? What are they interested in?

## Outcomes

Monitoring the outcomes you seek is an integral part of ROMA. You should have already set out the outcomes as part of the process of finalising your influencing objective, as outlined in Chapter 2. Refer back to Table 2, Step 2 in Chapter 2 for the discussion of nine possible outcomes to monitor and the different measures you can use to assess them.

## Context

External context is the final area to consider for monitoring. This is important to ensure the continued relevance of the interventions chosen. By this stage, you should be abreast of the shifting politics in your field of work: the agendas and motivations of different actors, who is influencing whom and any new opportunities for getting messages across. You should also be aware of new evidence emerging, or changing uses or perceptions of existing evidence, as well as the wider system of knowledge intermediaries, brokers and coalitions to use. Chapter 1 introduced three dimensions of complex policy contexts: distributed capacities, divergent goals and narratives and uncertain change pathways. Here are the areas to monitor for each.[12]

When distributed capacities define the context, it is helpful to monitor:

**1** **The decision-making spaces:** when, where, how are decisions being made? How are they linked?

**2** **The policy actors involved:** who are they? What are their agendas and motivations? How much influence do they have? Who are they influencing? How are they related formally or informally?

When divergent goals and narratives define the context, it is helpful to monitor:

**1** **Prevalent narratives:** what are the dominant narratives being used to define the problem? Who is pushing them and why? What opportunities do they offer?

**2** **Directions for change:** what are the different pathways already being taken to address the problem and (how) are they aligned to others?

When uncertain change pathways define the context, it is helpful to monitor:

**1** **Windows of opportunity:** are there any unexpected events or new ideas that can be capitalised on? Is there anything that can provide 'room for manoeuvre'?

12. Note the division between the three types of problem is not strict and it may be helpful to consider all five points.

**3**

From M&E to monitoring and learning

## How to use the measures

These individual measures should be treated as a menu from which to choose when developing an M&E plan. Table 9 is the key table to fill in:

- List those measures you are already monitoring (ask yourself: are you monitoring them in sufficient detail?)

- Identify another three or four measures you would like to monitor.

- Use this expanded list to populate the cells in Table 9, identifying which measures will help you meet each of the nine learning and accountability purposes. For example, the quality and quantity of outputs may be used to demonstrate financial accountability (you have spent the money on the outputs you said you would produce), improving operations (you produced them in a timely fashion) and deepening understanding (they represent a significant advance in your knowledge of the issue).

- Look across the table to identify any gaps; if the gaps are significant (i.e. the story of your intervention cannot be told properly), refer back to the list of measures above to work out how to fill them in. Choose the measures that match your intervention and the desired changes they are contributing to.

# 2. HOW TO MONITOR – COLLECTING AND MANAGING DATA

Part 2 looks at the different tools and approaches for collecting and managing the information needed. This is broken down into two types of method:

1. Methods to be used in real time for managers and practitioners to collect data throughout the process: these generally relate to output, uptake and more immediate outcome measures, as they tend to be more tangible and observable.

2. Methods more oriented towards the more intermediate and longer-term outcome measures: these require more time and are generally used retrospectively.

## Real-time data collection methods

Generally, if the intervention is very brief and engagement with individuals is very limited (e.g. through the broadcast media), the data for collection will be thin and may need to be supplemented with data from discrete studies. The deeper the engagement, the more in-depth will be the information you can collect in real time – and the more important these methods will become. Here are some of the methods.

### Journals and logs

One of the most basic ways of capturing information is by keeping a journal of observations, trends, quotes, reflections and other information. Logs are usually quantitative and simple – number of people attending an event or airtime during a radio show. Journals are more descriptive, and either structured with a specific format and fields to be filled in (such as progress against predefined measures or changes in contextual factors) or unstructured, allowing the author to record comments. They can be notebooks carried by team members or electronic (website, database, intranet, email or even mobile apps).

Examples include ODI's 'M&E log', which all staff members can contribute to by sending an email to a particular inbox, which then stores the information on the institute's intranet. The unstructured approach makes it very easy for staff to submit evidence of uptake of research outputs and feedback from audiences but does require effort to maintain, systematise and use.

Regular journals work well with small teams but become more challenging with larger teams. Oxfam GB introduced journals to its Climate Change Campaign Team as part of its M&E system to enable systematic documentation of the day-to-day monitoring that happens naturally. As the team expanded it became unwieldy to manage regular analysis of the journals, so the approach was modified. Staff now fill in the journals during their monthly team meetings.

The Accountability in Tanzania programme collects journals from its 20-plus partners, each reporting on the outcomes of up to 8 different actors, to understand their influence at national and local level in Tanzania. It asks for journals to be submitted only twice a year and has developed a database to organise the information, enable analysis and identify patterns.

### After action review

The US Army developed after action reviews as a technique for debriefing on a tactical manoeuvre. They have been adapted to organisational use and are commonly applied as part of a learning system. An after action review is typically used after an activity has taken place, bringing together the team to reflect on three simple questions: what was supposed to happen, what actually happened and why were there differences? They are designed to be quick and light – not requiring a facilitator, an agenda or too much time – and collect any information that might otherwise be forgotten and lost once the event passes. Therefore, they should be included as part of the activity itself and scheduled in right at the end. Like a journal, notes from the meeting should be filed away and brought out at the next reflection meeting.

A variation on the after action reviews is an 'intense period debrief', developed by the Innovation Network in the US as a method for advocacy evaluations. The richest moments for data collection in any policy-influencing intervention are likely to be the busiest – such as when mobilising inputs into a parliamentary committee hearing or responding to media attention. Data collection methods should adapt to this. The intense period debrief unpacks exactly what happened in that busy time, who was involved, what strategies were employed, how the intervention adapted and what the outcomes were, without interrupting the momentum of the intervention.

### Surveys

Surveys can be useful for obtaining stakeholder feedback, particularly when interventions have limited engagement with audiences. They are most appropriate for collecting data on uptake measures, since this is about reactions to and uses of intervention outputs. Surveys can also be used for outcome measures, but timing has to be considered, since outcomes take time to emerge. If a survey template is set up prior to an intervention, it can be relatively quick and easy to roll out after each event or engagement. This could be automated with an online service like SurveyMonkey – you just provide the link to your audiences.

### Web analytics

Since more and more interventions used in influencing policy are web-based, it is important to have a strategy for collecting information about use of web services: what is being seen, shared and downloaded, when and by whom. Website analytics are generally easy to set up, with services like Google Analytics providing free data collection and management.

Nick Scott at ODI offers good advice[13] on tracking a range of statistics, including webpages, publication downloads, search engine positioning, RSS feeds, Twitter, Facebook, mailing lists, blogs and media mentions. For each of these there are specific online tools recommended for collecting data. Once set up, these data services will run in the background and data can be collected and analysed when needed. Nick also describes the use of dashboards for compiling and visualising the data from multiple sources for analysis and use in decision-making.

Web analytics need to be used modestly and cautiously, however. They will never be able to replace the other data collection methods mentioned above; for example, they will never tell you exactly who is reading your work, who they work for, what their job is and what, if anything, they will do with it after they have read it.

13. http://onthinktanks.org/2012/01/06/monitoring-evaluating-research-communications-digital-tools/

**3**

From M&E to monitoring and learning

# Retrospective study methods

The real-time methods are unlikely to provide much data and insights at the outcome level. For this, you will need either to set aside time and undertake your own retrospective study or to commission a specialist to investigate for you. Either way, the following methods and approaches are useful to consider, as they are oriented towards the kinds of outcomes discussed in Chapter 2 and set out in Table 2.

### Stories of change

A story of change is a case study method that investigates the contribution of an intervention to specific outcomes. It does not report on activities and outputs but rather on the mechanisms and pathways by which the intervention was able to influence a particular change, such as a change in government policy, the establishment of a new programme or the enactment of new legislation. The change described can be an expected change that the intervention was targeting or an unexpected change – which itself can be positive or negative with respect to the original objective. Stories can also describe how an intervention has failed to influence an expected change, in which case they analyse the possible reasons why.

There are three major steps to writing a story of change:

**1** **Choosing the story:** this is usually prompted by the emergence of a success (or failure), through any of the data collection methods described above.

**2** **Gathering the evidence:** to really understand the contribution of the intervention and provide a plausible argument, you will most likely have to find additional information. This will involve interviewing key stakeholders and programme staff to trace the influence of your work and identify the mechanisms leading to change. This should also involve an element of substantiation of claims that the intervention has had an influence, for example by consulting experts in the field or those close to the change.

**3** **Writing the story:** stories should be relatively short (two to four pages), written as a narrative that is easy to read and leaves an impression. It should make a clear case for the intervention: describe the situation or the challenge it was responding to and how it intended to engage; explain who was doing what, when and to what effect; and discuss success or failure factors and any lessons for future interventions. Depending on the primary learning purpose, different emphases can be placed on different elements.

Stories of change are used in ODI as part of an annual review of the work of the institute. Researchers are encouraged to submit stories of the impact of their work to an annual competition, with the best published in the annual report and presented at the annual staff retreat. CAFOD, Tearfund, the Canadian International Development Agency and the UK Department for International Development (DFID) also use stories of change or case studies for understanding policy influence.

### Episode studies

Another case study method relates to episode studies, which look at the different mechanisms leading to a change. These are not systematic assessments of how much each factor has contributed to the change but they are very labour- and evidence-intensive. The steps are the same as for stories of change except that the evidence-gathering stage investigates any and all factors influencing the change, including but not limited to the intervention. This generally requires access to those close to the decision-making around the change in question. The advantage of this approach is that it can highlight the relative contribution of the intervention to the change in relation to other influencing factors and actors.

**Bellwether interviews**

The bellwether method was developed by the Harvard Family Research Project to determine where a policy issue or proposed change is positioned on the policy agenda, the perceptions of key actors and the level of traction it has among decision-makers. It involves interviewing influential people, or 'bellwethers', including elected representatives, public officials, the media, funders, researchers/think-tanks, the business community, civil society or advocates.

The method is similar to other structured interview techniques but with two important differences. First, at least half of the sample should have no special or direct link to the policy issue at hand. This will increase the likelihood that any knowledge will owe to the intervention rather than personal involvement. Second, bellwethers should be informed of the general purpose and topic of the interview but not be given specific details until the interview itself. This will ensure their responses are authentic and unprompted. The interview should start by being general and gradually become more specific.

---

## Box 12: Sample bellwether questions (from Coffman and Reid, 2007)

1. Currently, what three issues do you think are at the top of the [state/federal/local] policy agenda?

2. How familiar are you with [the policy of interest]?

3. What individuals, constituencies or groups do you see as the main advocates for [the policy]? Who do you see as the main opponents?

4. Considering the current educational, social and political context, do you think [the policy] should be adopted now or in the near future?

5. Looking ahead, how likely do you think it is that [the policy] will be adopted in the next five years?

6. If [the policy] is adopted, what issues do you think the state needs to be most concerned about related to its implementation?

7. Currently, what three issues do you think are at the top of the [state/federal/local] policy agenda?

8. How familiar are you with [the policy of interest]?

9. What individuals, constituencies or groups do you see as the main advocates for [the policy]? Who do you see as the main opponents?

10. Considering the current educational, social and political context, do you think [the policy] should be adopted now or in the near future?

11. Looking ahead, how likely do you think it is that [the policy] will be adopted in the next five years?

12. If [the policy] is adopted, what issues do you think the state needs to be most concerned about related to its implementation?

---

**System or relational mapping**

When the outcomes desired are related to how a system operates – for example building relationships between actors, shifting power dynamics, targeting the environment around which a policy is developed or improving information access or flows – then it can be useful to map that system to see how the different parts fit together. The data required for this are relational (i.e. to do with relationships, connections and interactions) rather than attributes (i.e. to do with facts, opinions, behaviour, attitude). They are usually collected through standard techniques such as surveys, interviews and secondary sources. By asking about the existence and nature of relationships between actors, a very different picture emerges of what the system looks like. This can be easily turned into a visual map to help identify patterns and new opportunities for influencing.

One particular method is NetMap, an interactive approach that allows interviewees to use physical objects and coloured pens to describe relationships between actors and their relative influence on a particular issue. It can be a useful variation if the aim is to gain perspectives across a system or network.

Another variation is influence mapping, which asks specifically about the influence one actor has on the opinions and actions of another. An influence map can show the primary and secondary (and if needed tertiary) influences on a key decision-maker. This can help in planning or adapting influencing strategies or identifying possible individuals to consult for a bellwether interview.

# 3. MAKING SENSE OF LEARNING AND DECISION-MAKING

Data need interpreting and making sense of. This part of the guide looks at the concept of sense-making and how monitoring systems can be fit for your purposes.

Sense-making is the process by which data are turned into actionable insights by subjecting them to beliefs and values, existing theory and other evidence. This can happen consciously through structured causal analysis with explicit parameters and questions. It also happens unconsciously through the social interactions and the periodic reflections that make up a natural working rhythm. Ideally, you need to be able to harness both for learning and accountability purposes.

## Making space for sense-making

Sense-making can take place in spaces and have particular rhythms. Spaces are the formal and informal meetings and events that make up the everyday life of organisations and programmes. Rhythms are the patterns and structures in time through which an organisation can direct, mobilise and regulate its efforts. Examples include annual reports, monthly team meetings, quarterly board meetings, end-of-project reports, independent evaluations, field trips, stakeholder consultations, phone calls with partners, weekly teleconferences, email discussions with peers and impromptu conversations. Each of these has different purposes (therefore different information needs); different rhythms and timing (therefore different levels of detail required); and different people involved (therefore different perspectives to draw on).

Sense-making can operate at the macro or micro level. The macro level relates to questions about strategy and the external context, looking at broad patterns and knowledge that can be applied elsewhere. At the micro level, the questions are about this particular intervention and these particular actors, and how to improve what is being done. Different spaces will have a different balance of micro and macro. Our monitoring system should make room for both kinds of sense-making in the appropriate spaces and maintain the balance between looking immediately ahead and looking to the horizon.

## Designing monitoring for sense-making

An effective monitoring system will identify the spaces and rhythms that already exist and weave them together with a common framework to provide structure for learning and reflection. ROMA provides such a framework; this is one of the major strengths of the approach. By walking a team through the development of a theory of change grounded in analysis of the context and system dynamics, ROMA provides a common language and schema that can be applied (explicitly or implicitly) within the sense-making spaces. For example, ROMA helps define the key stakeholders and the outcomes surrounding them that are important for the policy objectives. This can help greatly in identifying patterns from data by narrowing down where to look and what to look for.

It is important that sense-making is not constrained dogmatically to any framework, as this could mean important unanticipated changes are missed. Actions can have three broad effects: expected and predictable (e.g. you invite someone to a meeting and they turn up); expected and unpredictable (e.g. at some point after the meeting they remember what you were saying and recommend your work to a third party); and unexpected and unpredictable (e.g. that third party then shares your findings and claims credit for themselves). Your intervention will in reality engage with all three types of effect at the same time. Monitoring should be open to each of these, although it is usually the unpredictable effects that require the most attention from sense-making. The unexpected effects are trickiest to identify but often yield the richest learning. By structuring informal sense-making and employing formal causal analysis you can deal with these effects and ensure the right kind of data is supplied at the right time.

# Practices for informal sense-making

Informal sense-making happens all the time – we notice things, judge them, weigh them and assign value and significance to them. But it predominantly happens as a social process when interacting with colleagues or partners or struggling with a report. Monitoring can help make informal sense-making more systematic and conscious, and better linked to decision-making. The following practical tips can help make the most of these moments.

### Establish a common language

ROMA provides structure to sense-making by establishing monitoring priorities and signposting where to look for outcomes. The process of deciding key stakeholders to influence and developing the intended outcomes for each is extremely advantageous. It provides a common language for a team to use when making observations and weighing up the importance of information – and to know what you should be making sense of. It also provides a schema on which to base conversations – even, at a very practical level, an agenda for a review meeting. To enable quick and responsive sense-making, it can be good practice to frame outcomes in terms of key stakeholders' behaviour.

Example: consider stakeholder outcome measure 1 (see Table 2, Step 2, Chapter 2) 'attitudes of key stakeholders to get issues onto the agenda'. When using this measure for monitoring, you might create an indicator like 'governing party officials have positive attitude to tackling [issue]'. But how will you know their attitude has changed? A better indicator would describe what you would see if their attitudes were changing so when you see it you know it is significant. So you might use this instead: 'governing party officials request evidence on [issue]' or 'governing party officials make speeches in favour of tackling [issue]'.

Another practical tip is to ask these three questions for each piece of information collected:

- Does this **confirm** our expectations?

- Does it **challenge** our assumptions?

- Is it a complete **surprise**?

These questions will quickly help you decide what to do with that information. If it confirms, then use it as evidence to strengthen your argument, but it may indicate your monitoring is too narrow and you need to consider broader views. If it challenges, then review your assumptions and strategies and consider if the intervention is still appropriate. If it is a surprise, then take time to consider its implication and whether the context is changing.

### Draw on a variety of knowledge sources

To identify and understand unexpected effects, you need to be open to diverse sources of knowledge and not be constrained to a narrow view of what is happening. This means drawing on multiple sources of data but also diverse perspectives for making sense of these data. This might mean creating spaces that bring in different perspectives, for example inviting 'critical friends' into reflection meetings, discussing the data with beneficiaries to seek their input or searching for studies in other fields that shed light on what is going on.

One particular approach that specifically seeks to do this is the Learning Lab developed by the Institute of Development Studies. The Learning Lab is a three-hour structured meeting with participants invited because of common knowledge interests rather than common experiences (this means the meeting is not confined to a project team as such). The meeting is structured around four questions: what do we know, what do we suspect, what knowledge and practice already exist and what do we not know or want to explore further? An important aspect is a 20- to 30-minute silent reflection that allows all participants to think through what they already know or have heard or seen about the chosen topic. Participants are asked to draw on their practical experiences, including literature, discussions and observations.

### Use visual artefacts

Visualising data can greatly improve our ability to spot patterns and form judgements. But it can also take a lot of time and effort to produce meaningful graphs from a stack of data. Dashboards can help with this by automatically combining data from multiple sources and displaying it in a predefined way. Dashboards can be real-time, although this requires custom software or strong programming skills; or they can be produced on demand, but this requires more time and effort. ODI has developed a dashboard that can track all kinds of analytics in real time using a data aggregator called QlikView. The ODI Communications Dashboard brings together website and social media statistics, impact log entries, media hits and more in one visual report than can be filtered by output or programme.

A simpler alternative to a dashboard is a traffic light system to alert you to events that require attention. For example, you might use a stakeholder database to track information on voting patterns in parliament. A quick way to make sense of this information would be to assign a colour to each stakeholder depending on how they are voting (green = as expected, yellow = unexpected but probably doesn't matter, red = unexpected and will affect our programme).

Relational data, such as who has been meeting with whom or who has shared your work with whom, can be visualised on a network map. This will allow for filtering and clustering of relationships to uncover patterns and understand the dynamics of policy communities.

# Practices for formal causal analysis

Some instances may require a more structured enquiry than informal sense-making. Formal causal analysis is a critical part of a learning system and can be conducted internally or externally (e.g. through assessments, reviews, evaluations or research studies). Practitioners and evaluators alike can use a few techniques for policy-influencing interventions.

### Compare with the theory: process tracing

One of the most plausible ways to understand causes in complex contexts is to compare observations with a postulated theory.[14] For example, our theory says if we train junior parliamentary researchers to interpret and use scientific evidence they will use these skills to better advise parliamentary committees, which in turn will draft more appropriate and effective bills. In this example, each stage can be tested by comparing the data on what the researchers did after their training, and the subsequent decisions of the committees they are working with, with the effect expected. It is then possible to confirm or rule out particular causal claims.

This is the basis of process tracing, a qualitative research approach used to investigate causal inference. Process tracing focuses on one or a small number of outcomes (possibly involving a process of prioritisation to choose the important ones) to verify they have been realised (e.g. a policy-maker makes a decision in line with recommendations). It then applies a number of methods to unpack the steps by which the intervention may have influenced the outcome. It uses clues or 'causal process observations' to weigh up possible alternative explanations. There are four ways clues can be tested:

**'Straw-in-the-wind' tests:** when a straw seems to be moving, it lends weight to the hypothesis that there is wind but it does not definitively rule it in or out (e.g. we sent our report to the policy-maker but do not know if it got to them).

**'Hoop' test:** a hypothesis is ruled out if it fails a test (e.g. was the report sent to them before the decision was made?)

**'Smoking gun' test:** seeing a smoking gun lends credence to the hypothesis that it was used in a crime but is not definitive by itself (e.g. we see our report on the desk of the policy-maker but don't know if they have read it).

**'Double decisive' test:** where the clue is both necessary and sufficient to support the hypothesis (e.g. we observe precisely the same language in the decision of the policy-maker as in our recommendations).

---

14. Step 3, Chapter 2 describes steps for developing a theory of change.

## Check timing of outcomes

A strategy that can help determine causal inference is laying out all the outcomes in a timeline to demonstrate the chronology of events. If you also place the intervention activities and outputs on the timeline then you can begin to establish causal linkages, visually applying the test that effect has to follow cause. This can eliminate many competing claims about causal inference and help narrow down the important ones. There may also be timings inherent in the theory of change, so this can also be used to judge the plausibility of contribution.

The RAPID Outcome Assessment approach has been used to determine the contribution of research to policy change. In RAPID Outcome Assessment, a timeline is mapped of milestone changes among pre-determined target stakeholders alongside project activities and other significant events in the context of the intervention. A workshop is convened with people close to the intervention and the changes described. Participants work through each of the changes observed and use their knowledge and experience to propose the factors that influenced them (which could be the intervention or other factors) and draw lines between the different elements of the timeline.

Figure 8 is a timeline developed to analyse the Smallholder Dairy Project in Kenya, a research and development project that aimed to use findings to influence policy-makers. The analysis, based on the ROA approach, identified the key actors setting and influencing policies affecting the dairy sector in Kenya. It then used interviews to find significant behavioural changes in these actors, triangulating them in a stakeholder workshop. This resulted in a set of linkages between the changes and key project milestones as well as external events.
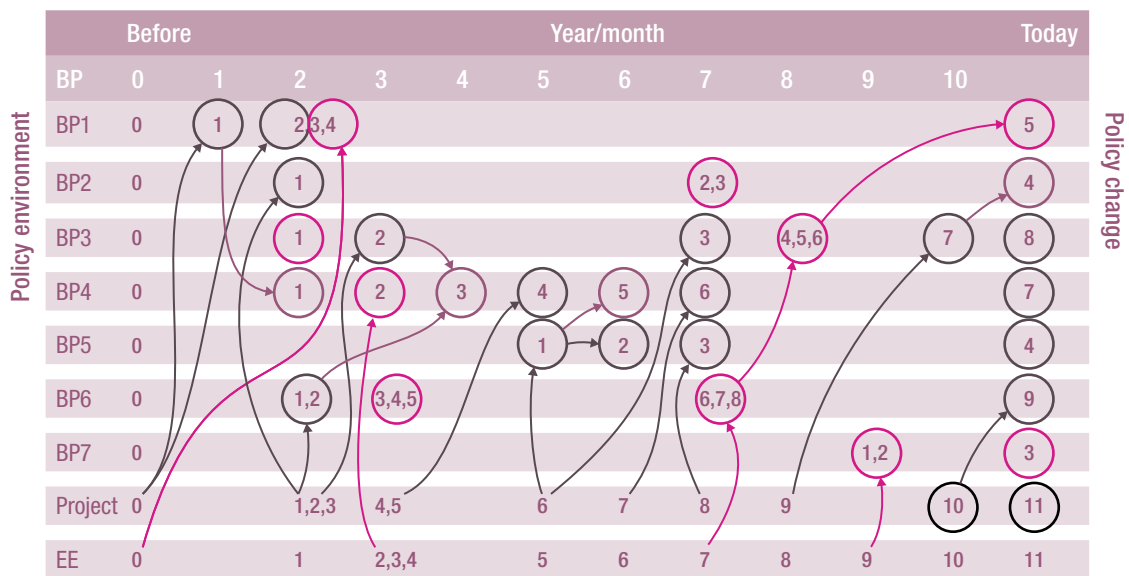


**Figure 8: Example of timeline showing changes observed in seven key stakeholders (BP1-7), project milestones and external environment (EE) for Smallholder Dairy Project, Kenya (from ODI, 2012)**

### Investigate possible alternative explanations

All the approaches described above share a commonality: they all look beyond the intervention for possible contributing factors. It is fairly obvious that, when working in open systems, we are rarely the sole actor trying to influence outcomes. It is vital, then, that whatever contribution is made is placed in the context of all the other actors and factors operating in the same space. Investigating alternative explanations can help gauge the relative importance of the intervention, but can also help narrow down hypotheses to test if alternatives can be ruled out – which can strengthen the case for intervention contribution.

This is the basis of most theory-based evaluation approaches. It is the core purpose of the General Elimination Methodology developed by Michael Scriven, which systematically identifies and tests 'lists of possible causes' for an observed result of interest. As well as collecting data about the intervention, the study collects data about other possible influences so as to either confirm or rule them out.

The General Elimination Methodology was used in an evaluation of a public education campaign to end the juvenile death penalty in the US. The campaign, funded with $2 million from a collaboration of foundations, ran for nine months from 2004 to 2005 during a US Supreme Court hearing to review a number of cases of juvenile offenders facing the death penalty. On 1 March 2005, the Supreme Court ruled that juvenile death penalties were unconstitutional. The evaluation sought to determine to what extent the campaign influenced this decision.

Following the General Elimination Methodology, the evaluation started with two primary *alternative* explanations: 1) that Supreme Court justices make their decisions entirely on the basis of the law and their prior dispositions rather than being influenced by external influences; and 2) that external influences other than the final push campaign had more impact.

The evaluation gathered evidence through 45 interviews, detailed review of hundreds of court arguments and decisions and legal briefs, analysis of more than 20 scholarly publications and books about the Supreme Court, news analysis, reports and documents describing related cases, legislative activity and policy issues and the documentation of the campaign itself, including three binders of media clips from campaign files. Through all this evidence the evaluators were able to eliminate sufficiently and systematically the alternative explanations to arrive at an evidence-based, independent and reasonable judgement that the campaign did indeed have a significant influence on the Supreme Court decision.

# 4. SUMMARY

This chapter started out with the aim of providing practical advice for people working to influence policy to build reflective and evaluative practice into their work to support decision-making and demonstrate progress.

Part 1 introduced nine 'learning purposes' – the overarching reasons for undertaking any kind of M&E activity that should drive the design and use of M&E. It proposed 35 individual measures for policy-influencing interventions across six categories (strategy, management, outputs, uptake, outcomes and context), and suggested how these could be used for the learning purposes.

Part 2 discussed how data could be collected both in real time, as the intervention is being carried out, and in retrospect, through detailed studies.

Finally, Part 3 turned to the important task of making sense of those data and putting them to use in decision-making and demonstrating impact.

Since the theme of the chapter has been evaluative practice, it is apt to conclude with a few final pointers on good practice:

**1** Put use at the heart of your monitoring, evaluation and learning to make sure any enquiry will have a positive contribution.

**2** Be grounded in theory from the beginning and test each stage as you go.

**3** Consider competing theories so as not to close down unintended effects.

**4** Embrace failure as just as good an opportunity to learn from as success.

**5** Invest in your monitoring and learning in proportion to the scale of your intervention: sometimes it is appropriate to use simple measures.

**6** Be conscious of rhythms and spaces in which learning occurs: it happens at different paces and different levels.

Finally, there is a traditional African proverb that encapsulates the attitude to take when developing M&E systems for policy influence: 'we make our path by walking it'. Start by looking at what people are already doing, where data are already collected and the spaces that already exist for sense-making, and then work to strengthen and support those. If existing patterns are ignored, efforts may be wasted because people will always drift towards the familiar and the easy.