

There are specific challenges to address when monitoring and evaluating capacity building interventions. Planning and preparation are important, as is a recognition that any monitoring and evaluation (M&E) processes need to support rather than inhibit capacity building. Many different tools and methodologies are available, but these will often need to be adapted to the local context. Evaluating progress across a portfolio of organisations is difficult and involves trade-offs between consistency and local ownership.

Capacity development is generally recognised as a deliberate process whereby people, organisations or society as a whole create, strengthen and maintain capacity. Capacity building is understood as a purposeful, external intervention to strengthen capacity over time (see Simister and Smith 2010).

Capacity building may be relatively straightforward to monitor and evaluate if a provider is only working with a relatively small number of organisations. But the challenges may be multiplied when working with a larger portfolio. In these cases, capacity building providers often fall into one of three traps (see James 2009):

- doing nothing, paralysed by the difficulties and costs;
- setting up a mechanical or misleading system;
- setting up something that is too complex or burdensome on the organisations they are supposed to be supporting.

However, INTRAC's experience is that capacity building can be effectively monitored and evaluated with an appropriate blend of methodologies, time and patience. This will generally require effort, resources and good planning, but the rewards in terms of enhanced capacity development are worthwhile.

This short paper first looks at some of the theoretical issues surrounding the M&E of capacity building. It then lists some of the different tools and methodologies that are available. Finally it addresses some of the challenges faced when assessing and summarising the results of capacity building across a large portfolio of work.

PART A: Theoretical Issues

Challenges: There are a number of challenges associated with the M&E of capacity building work that are not always present when assessing more straightforward service delivery work.

- Capacity is an intangible and sometimes contested concept.
- It can take a long period of time for capacity building interventions to result in enhanced capacity, or to filter through to ways of working.

- Results may be spread across many organisations, and M&E work may need to be coordinated across long chains of actors including donors, capacity building providers, recipients and intended beneficiaries.
- Capacity development is not a linear process. Organisations and individuals evolve over time in response to changing internal and external environments, and it can be hard to separate out these changes from more purposeful, intended ones.
- Most methodologies for monitoring and evaluating capacity development require some level of self-evaluation. But individuals within organisations may feel unable to provide honest and open opinions, especially if these are taken as criticisms of working colleagues, or if results are linked to funding decisions.
- The central purpose of capacity building should be to enhance capacity. Any M&E work that undermines this process is actually working against the capacity process itself.

Case study: Evaluation of MFS II

Between 2012 and 2014, INTRAC worked on an evaluation of Bangladeshi agencies supported through the Dutch Ministry of Foreign Affairs, via several NGO consortiums. As part of the work, INTRAC designed a self-evaluation process for the Bangladeshi agencies to assess their capacity at intervals in order to establish change. The process was designed and conducted in a participatory way.

It was later revealed that one of the Bangladeshi agencies concerned had previously been the subject of a number of different organisational assessments carried out by different supporting agencies. In at least one case the consultant facilitating the organisational assessment had independently 'marked' the organisation concerned, and had not told staff the results. INTRAC believes this kind of assessment is detrimental to the capacity development process itself.

Setting the Framework: Any type of M&E work relies to a large extent on understanding what an intervention is attempting to do, and what changes it is attempting to bring about (and why). In capacity building interventions this is especially important. There are three critical areas to clarify: the purpose of the capacity building; the

purpose of M&E; and the nature of accountability between a capacity building provider and other agencies.

Establishing the purpose of a capacity building intervention may be the first step. At an organisational level, capacity building can be divided into two types. *Technical* capacity building attempts to address a specific issue, such as an organisation's ability to develop funding proposals, manage health centres or teach pupils. Technical capacity building is often carried out within a specific project or programme, in response to clearly stated needs. It is therefore relatively easy to address the 'capacity for what' question.

On the other hand *organisational* or *programmatic* capacity building is provided to support organisations to fulfil their core functions and achieve their own mission. This type of capacity development may be slow and complex, requiring in-depth reflection on an organisation's culture, values and mission. The goal of this kind of work is to help improve an organisation's overall performance and ability to adapt itself within a changing context. In such cases it is important to address some of the key issues that are nowadays commonly addressed through Theory of Change thinking:

- Why is the capacity building work being done, and why now?
- How is the capacity change expected to occur?
- How is individual or organisational change expected to contribute to wider change?
- Why is this wider change important?
- How will this fit in with the work of others?
- What are the key assumptions behind this work?

The next step is to establish the purpose of the M&E system. Many systems are set up with the intention of providing a mixture of learning (in order to improve performance) and accountability to donors or supporters. But these purposes may be in conflict, and there may be significant differences in the type of information collected, the methods used to collect it and the integrity with which M&E is carried out. One issue that INTRAC has often noted is that organisations primarily basing M&E on their need to be accountable to donors are much more likely to take (positive) results at face value, and corresponding less likely to probe behind these results for deeper, hidden meanings.

It is also important to recognise that there may well be competing demands on M&E across different organisations from donors, through to capacity building providers and recipients, and then to partners and beneficiaries. In many cases the challenge is to reconcile these competing demands.

Finally, the nature of accountability needs to be understood. In many cases the primary accountability is (or should be) between a capacity building provider and the recipient organisation. But more often accountability is skewed towards those providing the funding. In these

cases it is important to establish exactly what an M&E system needs to measure in order for a capacity building provider to demonstrate accountability.

At the most basic level, capacity building providers can be held accountable for activities and outputs. This includes accounting for money spent and trying to ensure that any work carried out is both the right thing to do and done as well as possible. Capacity building providers can also be held accountable through their outcomes, which are generally assumed to be the changes within the recipient organisations. Some feel this is unfair as a capacity building provider cannot control capacity within another organisation, although most believe it is reasonable to expect capacity building providers to at least report on initial changes arising out of their work, whether positive or negative.

More controversial is the view that capacity building providers can be held accountable for wider changes resulting from enhanced capacity within supported organisations, such as changes in beneficiaries' lives. Yet this is a pressure that many capacity building providers are currently experiencing.

Deciding how far to measure: In response to this pressure, capacity building providers need to decide how far their M&E systems should attempt to identify ultimate changes in beneficiaries' lives. Of course, to some extent, this depends on the purpose of the capacity building support. For example, it is much easier to assess the ultimate results of capacity building work aimed at enhancing medical practice with patients than work aimed at changing the culture of a large, organisation working across multiple sectors and countries.

In reality, in the latter case it is rarely possible to really *measure* the change brought about through general or programmatic capacity development. But it is often possible to *illustrate* some of the changes resulting from enhanced capacity by developing stories of change or case studies linking change to beneficiaries back to the provision of capacity building support.

It is also important not to forget the process itself. Capacity building providers need to be honest and open enough to seriously monitor and evaluate their processes. This should include giving the recipients of capacity building support opportunities to say how well (or how badly) they think that support was provided.

PART B: Tools and methodologies

There are many circumstances where changes in capacity can be measured directly, particularly where technical capacity building is concerned. For example, changes in fundraising capacity can be measured by recording changes in the number of external funders supporting an organisation, or the amount of revenue generated. Equally changes in the capacity of organisations to support more inclusive teaching methods can be assessed by measuring changes in exam results. Effectively, direct measurements such as

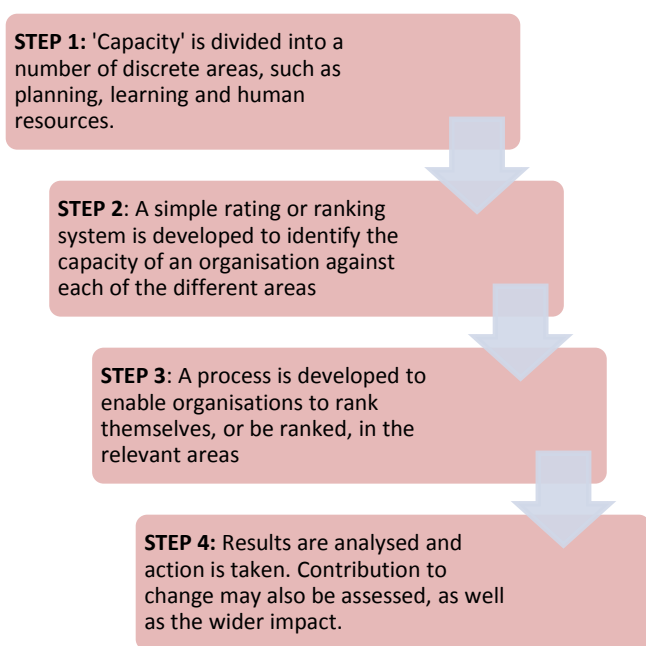
this bypass the 'capacity' measurement problem itself and go directly to the end result. (It is easier to measure whether a local administration is clearing refuse than it is to assess its capacity to do so).

But in cases where the end results are not so clear cut there are a variety of different tools and methodologies that are available. This section of the paper lists some of them.

Organisational Assessment (OA) tools: Often known as Organisational Capacity Assessment Tools (OCATs), these tools are perhaps the only M&E tool or methodology in widespread use designed specifically with capacity building in mind. OA tools can be used in three distinct ways:

- to assess the capacity of an organisation to act as a partner or recipient of funds;
- to make a general assessment of organisational strengths and weaknesses as part of a needs assessment; and
- to help monitor and evaluate progress by showing change over time.

There are many different types of OA tools available. Most, however, are based on a similar pattern of steps.



OA tools can be used for M&E in two ways. Firstly, a needs assessment can lead to the development of an action plan with associated objectives and indicators – perhaps in a logical framework or similar results framework – and this plan can be monitored over time. Secondly, an organisational assessment can be repeated at discrete intervals. Changes in scores are then used to show how capacity has changed within an organisation. If necessary, these changes can also be investigated to assess whether or how far they are the result of a particular capacity building intervention.

There are many different views about OA tools, and their strengths and weaknesses. Some of these are as follows (see Simister and Smith 2010).

Strengths	Weaknesses
<ul style="list-style-type: none"> • They ensure that capacity building is formally monitored and evaluated • They enable organisations to identify necessary changes to help achieve their mission • They provide a rolling baseline so that progress over time can be assessed • Results can sometimes be aggregated or summarised across different organisations, sectors or countries • OA tools focus on the outcomes of capacity building work, not just the activities carried out • Unintended or negative consequences of capacity building work are also covered 	<ul style="list-style-type: none"> • It can be hard to show how improved capacity is attributable to any particular support provided • An OA tool does not necessarily show how any improved capacity contributes towards improved performance • Ranking or rating can be subjective, based on perceptions of different stakeholders • A lower ranking score does not always indicate weak capacity – it may be an indication of enhanced awareness of limitations • A higher ranking score may be the result of over-confidence in an organisation's capacities

Perhaps the biggest concern over the use of OA tools is that they encourage a blueprint approach. This means that Civil Society Organisations (CSOs) based in the South are expected to conform to the standards of an idealised, Northern, non-governmental organisation (NGO). Indeed, some organisations have been deeply critical of the practice of CSOs in the South “... being assessed against templates, checklists and models of a “best-practice” organisation developed in the North and having their capacity built accordingly” (Barefoot Collective 2009, p14).

However, in INTRAC’s experience there is a big difference between ‘normative’ OA tools and those designed more around functionality. A normative tool usually assesses an organisation against pre-defined criteria that may involve assumptions about what a ‘good’ organisation should look like. For example, an NGO may be rated more highly if it has a strategic plan or a set of core indicators, and regularly undertakes external evaluation.

On the other hand a more functional organisational assessment tool tries to assess whether an organisation can plan effectively, or carry out effective M&E, without making assumptions about how this might be achieved. These tools are based on the assumption that all organisations need to plan, exist, adapt, and serve their

core missions, but the way in which they do so may vary widely from organisation to organisation.

Perhaps the best known of the newer OA approaches is the Five Capabilities model, designed through the European Centre for Development Policy Management (ECDPM) project, and very popular with Northern European donors and NGOs. It identified five core capabilities, which, it is argued, if developed and integrated successfully, will contribute to the overall capacity of an organisation. The model of five capabilities is designed to provide a basis for assessing the capacity of an organisation and tracking it over time. The capabilities are (see Engel et al. 2007):

- to survive and act;
- to achieve development results;
- to relate;
- to adapt and self-renew; and
- to achieve coherence.

This model can theoretically be applied by a wide range of organisations stretching from large, International

NGOs through to organisations that exist for just a few weeks every year around a specific event such as World Toilet Day. It therefore helps to remove some of the problems associated with a blueprint approach to organisational assessment.

Scorecards: Some people use ‘scorecards’ to describe OA tools that are narrower than the holistic tools use to assess entire organisations. But actually the principles are the same – divide work into discrete areas, rank or rate capacity, take action on the findings – and in many cases there is little difference. In general, however, scorecards are designed to work across a narrower set of areas of capacity; commonly those areas that are being supported by capacity building.

For example, in an ongoing programme of Civil Society Support in Ethiopia (CSSP) INTRAC helped design a set of scorecards to be used by a range of different Ethiopian CSOs. The scorecards covered the main areas in which CSSP was providing support: including financial management; leadership and governance, project cycle management and engaging with core constituencies. One such scorecard is shown below.

Scorecard Area: Capacity (and commitment) to work with and for the poorest women, men, girls and boys

Low capacity	1	2	3	4	5	High capacity
<ul style="list-style-type: none"> • The organisation does not consult the people it claims to work with • The organisation has very minimal understanding of the different social groups and social structures • The organisation does not identify the different priorities determined by different groups of poor women, men, girls and boys • Poor, very poor and the poorest people have no role in evaluating the organisation’s work in the community 	<p>Score (Please mark one box only)</p>					<ul style="list-style-type: none"> • The organisation consults regularly with the people it claims to work with – particularly with the poorest or those hardest to reach. • The organisation has good understanding of the different social groups and social structures • The organisation includes poor or hard to reach people on its board. • The organisation adjusts its priorities, spending and staffing based on feedback from the poorest girls and boys, women and men • Consultations are arranged so different social groups (e.g. girls, people with disabilities) have separate opportunities to share their points of view • The poorest people have a role in evaluating the organisation’s work in the community

Initially, the M&E system was based around self-evaluation of supported CSOs, theoretically backed-up through facilitation from project staff. A set of results was achieved and analysed from over 100 supported CSOs. In fact, the results were very positive in terms of showing significant changes in capacity across all the different scorecard areas. However, on closer inspection many of the results were found to be unreliable. This was not through any deliberate falsification of results, but rather through lack of capacity of project staff to facilitate the process, lack of resources, and a natural bias on behalf of supported organisations to claim improvement.

Halfway through the programme CSSP decided to radically change the way it evaluated capacity change. With the agreement of its donors it decided to have all scorecards facilitated by a small core of dedicated staff who could apply the scorecards consistently across a range of different, sampled, organisations. In addition, the scorecards themselves have gone through many revisions, and CSSP has constantly striven to adapt and amend the process in the light of evolving understanding. Only now, after four years, does CSSP feel the information coming through the scorecard system is reliable enough to feed into management decision-making.

Outcome Mapping: Although not specifically designed with capacity development in mind, Outcome Mapping (OM) is regarded as an effective method of planning and reporting on capacity development work – much more so than the logical framework approach, which relies on predicted changes across set timescales. This is for a number of reasons:

- OM requires a project or programme to identify boundary partners – individuals, groups or organisations with which it interacts directly to effect change. OM is therefore particularly appropriate when assessing change at an organisational level (Earl, et. al. 2001).
- OM encourages a spread of possible outcomes (known as progress markers) ranging from initial changes one would expect to see over the course of a project or programme to changes one would like or love to see. This avoids the need for precise prediction of the pace of change, or reliance on any one indicator.
- OM focuses on behavioural change, and progress markers are designed to describe observable changes in actions, behaviours and relationships that are (or should be) straightforward to measure.
- OM deliberately recognises complexity, and the fact that capacity building providers are not ultimately responsible for changes within boundary partners (ibid).

The idea of Outcome Mapping is to set a series of progress markers and then to collect and analyse information at regular intervals, mapping this information onto the Outcome Map. Normally, Outcome Maps are developed independently for each supported organisation. However, the methodology also allows for common maps to be applied across a range of organisations if they are of similar nature.

An example of this is shown in the case study on the right (see MacDonald 2015). Again, this case study shows how an M&E system for assessing capacity development may need to be revised and refined over a period of some years if it is to remain useful

Stories of change: These are often used as an alternative to more numeric methods as they are perceived to be better capable of describing the richness and complexity of individual, organisational and societal change. However, unless an organisation is clear about how stories are generated and used, they can be dismissed as anecdotal. In response, a number of different methodologies are used to help introduce more rigour into the process.

Most significant change (MSC) methodology is often mentioned as an alternative to results-based management techniques. MSC is a system designed to record and analyse change in projects or programmes where it is not possible to precisely predict changes beforehand, and is therefore difficult to set pre-defined indicators. It is also a tool designed to ensure that the process of analysing and recording change is as participatory as possible.

Case study: Outcome Mapping in CDKN

The Climate and Development Knowledge Network (CDKN) supports negotiating groups from vulnerable and developing countries in climate change talks. Initially, a large Outcome Map was developed containing more than forty markers of change. Over the past four years these have gradually been whittled down to a core set of twelve (proposed) markers that can be applied across each of the organisations supported. Many of the markers point to changes in capacity directly, but many are the observable behavioural changes that should result from enhanced capacity. Note that capacity support within this programme contains a range of inputs such as technical assistance, resourcing, research and mentoring.

The OM in its current form is as follows.

Love to see markers:

- Final negotiation texts for international climate change negotiations include submissions from the poorest and most climate vulnerable countries and/or include the outcomes they prefer.
- Groups/countries/constituencies are asked to enter formal links with other (influential) groups
- There are a high number of "joint submissions" or "joint press conferences" made by groups/countries/constituencies or collaborations among groups / progressive countries, relating to key technical issues and negotiating tracks

Like to see markers:

- Delegations join appropriate groups or form cross-group coalitions based on shared progressive interests during international climate change negotiations.
- Groups / countries/ constituencies achieve increasing media coverage of their issues and demands
- Groups / countries increasingly identify and agree priorities or desired outcomes in advance of meetings within international climate change negotiations
- Delegates cite relevant legal precedents or technical research to support their positions or to challenge the wording in agreements
- Groups/ countries/ constituencies develop knowledge management systems that allow institutional memory to be captured and that support, for example, the rotation of roles such as Chair, or the tracing of developments in negotiating tracks over time
- Increased proportion of delegates have technical background and/or have been selected to attend meetings due to their technical background rather than their seniority

Expect to see markers

- Delegates make a greater number / proportion of interventions and submissions in areas relevant to their national or group interests.
- Groups/ countries/ constituencies increasingly access available advice and support during Conference of Parties (COPs) and other major conferences
- Delegates increasingly understand the technical and political issues behind the negotiations

The purpose of MSC is to identify significant changes brought about by a project or programme, especially qualitative changes that cannot easily be represented through numbers. MSC relies on people at all stages of a project or programme sitting together to identify what they consider to be the most significant changes within pre-defined areas (or domains). Although not specifically designed for capacity building programmes, MSC can easily be adapted for such purposes by defining a domain or domains around 'organisational capacity change'. For example, CCDB in Bangladesh created a domain around the sustainability of people's institutions, whereas MS Denmark included a domain on organisational performance (see *Davies and Dart 2005*).

One of the main points about MSC is that it involves a transparent process for the generation of stories that explains why and how each story was chosen. It is specifically not designed to produce representative stories. Instead it is designed around purposive sampling to find the most interesting or revealing stories.

Case study: CABUNGO

CABUNGO, a Malawian based organisation, used MSC to evaluate its capacity building services as a pilot project. The pilot enabled CABUNGO to identify changes in organisational capacity such as shifts in attitudes, skills, knowledge and behaviour. Changes were also seen in relationships and power dynamics. Most of the stories generated described internal changes within the recipient organisation, but some also described changes in their external relationships with donors and the wider community.

Participants in the evaluation process felt that the story-based approach was useful in helping CABUNGO understand the impact it had on the organisational capacity of its clients, and how its services could be improved. The key advantages of using MSC were its ability to capture and consolidate the different perspectives of stakeholders, to aid understanding and conceptualisation of complex change, and to enhance organisational learning. The constraints lay in meeting the needs of externally driven evaluation processes and dealing with subjectivity and bias (Wrigley 2006).

An alternative is to provide stories based on random sampling - choosing a selection of individuals or organisations as a focus for in-depth case studies. This then allows some extrapolation of findings which enable an estimation of the overall effects of a capacity building programme. However, significant resources may be required to generate enough information to draw wider conclusions.

Support to individual organisations can also be assessed using purely qualitative techniques. This involves developing a qualitative baseline (a story of what the situation is at the start of the support) and

describing a picture of what the situation might be in the future. Regular monitoring then builds a series of pictures over time, showing what has changed and why. For example, the CSSP project in Ethiopia has developed a journalistic approach to the development of stories, in which stories are not only identified, but also checked and probed for accuracy before a final narrative is developed.

In all these cases, stories of change (or narratives or case studies) can be used to develop a portfolio of changes and lessons that can be used to summarise the work of capacity building providers and/or capacity changes in supported organisations. Such stories can often be used for accountability and communications purposes.

But if they are truly to be used for learning and improving – or to withstand external scrutiny – there are two main features that need to be in place. Firstly, the methodology for selecting the stories must be transparent (there is nothing wrong with cherry picking provided you are open about the fact that you are cherry picking). Secondly, the stories themselves must be based on robust information that has been questioned and probed, through whatever method.

Tools of the trade: There are many standard tools that are used for M&E. These were not designed with capacity building in mind, but most can be applied to capacity building endeavours. Amongst the more straightforward tools are individual or group interviews, focus-group discussions, questionnaires, observation, use of diaries and timelines. More complex tools might include process tracing, outcome harvesting, tracer studies, appreciative inquiry and many more. All of these tools may be used in conjunction with the other methods described in this section.

One method that is often used to generate M&E information is a survey. These can be very useful, especially after training courses or other forms of capacity building events to assess immediate reactions. They can also be used to show perceptions of partners that are recipients of capacity building activities.

However, there are some dangers associated with using surveys to assess capacity change. With the advent of resources such as Survey Monkey a certain amount of 'survey fatigue' has set in, and response rates for surveys may be very low. This has been INTRAC's experience over the past few years. Where response rates are low, information can still be generated for communications purposes, and stories of interest may arise. But such information should not be treated as representative. There is a large element of self-selection associated with survey response. If providers of capacity building support wish to use surveys to assess change in a more methodological manner then they also need to take steps to ensure response rates are high.

Client satisfaction: M&E might also be based around client satisfaction. Sometimes, this can be established through interviews or surveys (see above) although questions of bias will always intrude where people are

expected to comment on those providing support. However, even if results are suspect the methodologies, if used sensitively and appropriately, can be seen as useful downwards accountability mechanisms.

Sometimes, capacity building is linked to other forms of support. For example, many Northern NGOs provide support to Southern NGOs as a condition of funding. But sometimes organisations purchase capacity building services from providers with money, or voluntarily devote staff time and resources to capacity development. In these cases M&E can be based around a more market-led approach, on the assumption that if organisations come back for more support (or persuade others to do so) it must be because they valued the previous support and found it useful. This can be seen as a valid M&E approach for demand-led capacity building work.

Another proxy measure of client satisfaction might be the extent to which capacity building resources are accessed. Some organisations monitor downloads of resources or track how often websites or blogs are accessed to gauge whether and how they are meeting the needs of different stakeholders.

Sense-making: All the methods described in this section can be used for generating information. But there are also many different processes which can be used to share information and to jointly make sense of key questions, such as what has changed, what was the contribution of the capacity building provider, what do the changes mean, why are the changes important, and what should be done differently in the future?

These include conferences, workshops, discussion papers, research studies, reviews, formal evaluations and impact assessments. Through carrying out a range of different exercises, different stakeholders can be brought together jointly to build up a picture of change, and make recommendations for the future. These exercises may also be useful in addressing wider aspects of capacity building such as the enabling or constraining environment, relationships, power dynamics, and wider impacts on targeted communities (see Lipson and Hunt 2008).

Triangulation: There has been recognition for some time that the results of capacity building work need to be assessed through a range of different tools, methodologies and approaches.

Some organisations combine traditional planning models, such as the logical framework, with more complexity-oriented methodologies such as outcome mapping or MSC. Others combine regular organisational assessments with periodic reviews or evaluations. Many use different methodologies to gauge the opinions of a variety of different stakeholders throughout the chain of support from donors to communities. In theory, at least, there are enough different tools and approaches to enable any organisation with sufficient commitment (and resources) to build up a picture of change, even in a complex area such as capacity development.

However, organisations also need to balance M&E requirements against the requirements of other functions and stakeholders. The challenge is often not so much how to conduct appropriate M&E from a technical point of view, but more about how to keep M&E systems light and flexible so that they do not impose unnecessary burdens on providers or recipients of capacity building support. In essence, the more that M&E can be built into a project or programme as a vehicle for capacity development itself, the more organisations will find it easier to justify the time and expense.

PART C: Summarising progress across a portfolio

Aggregation and summarisation: If facilitation of capacity building is well designed and implemented, it is not really that difficult to monitor and evaluate a single capacity building intervention. One or more of the different tools and methodologies outlined in the previous section can be selected with the active involvement of the recipient organisation, and these can feed into reporting, learning and decision-making.

Indeed it could be argued that if a capacity building provider is doing their job correctly no formal tool or methodology is needed at all. This is because a good provider should know what the needs of an organisation (or set of individuals) are. They should know how an organisation is reacting to different methods of a support. They should know what is changing within an organisation and why, and what needs to change further. In essence, a good capacity building provider, working alongside a recipient organisation in a participatory way, should be undertaking good monitoring as part of the capacity building process itself, and the two functions of capacity building and M&E should be interlinked.

However, where capacity builders are expected to monitor and evaluate a portfolio of work across numerous different organisations the challenges are much greater. Essentially it is much easier to summarise and aggregate performance in any field of work across a range of interventions where:

- measures of performance are clearly laid out and consistently applied;
- a common tool or methodology is used;
- work is carried out over similar timescales;
- the quality of information collection and analysis is consistent; and
- contributions to change are similar (or can be disaggregated) in the aggregation process.

Where capacity building is concerned, the issue is more pertinent because of the nature of the work. Ideally, a capacity building provider would want to select M&E tools and methodologies with the recipient of support, and possibly adapt these to local circumstances as well. But across a portfolio of work this may not be possible, and in order to summarise change across a portfolio a capacity building provider may need to impose the use

of common tools. At best this may be seen as inhibiting local ownership of the process, and at worst it might be detrimental to the capacity development process itself. But in many situations this may be an unavoidable task.

Of course it may not be necessary to measure change in every single supported organisation using a common tool, and in some cases a sample of organisations may be sufficient to generalise findings across a portfolio.

Indicators at portfolio level: Many organisations providing capacity building support are asked at some stage to develop objectives and indicators – either inside or outside of logical frameworks – to assess progress against a portfolio of work. At the output level this can be relatively straightforward as most providers can easily count how many organisations or individuals they support. It may also be straightforward to count some forms of capacity building support, such as

training, hosting events, access to resources, etc. although forms of capacity building support such as mentoring and partnership may be harder to quantify. Most organisations also find themselves able to assess the quality of their support with a judicious mixture of quantitative and qualitative post-event reaction measures.

But developing outcome indicators at portfolio level is often more of a challenge. The table below shows how this can be done, applying some of the methodologies described in part B. It is important to note that in these cases the standard order of play is reversed: rather than developing an indicator and then identifying a tool to collect it, the tool is identified first and the indicator is then developed from the tool. Most of these indicators can also be used to generate milestones and targets as well.

Method	Possible indicators	Notes
No consistent method	<ul style="list-style-type: none"> # of organisations with enhanced capacity # and description of capacity changes observed # of organisations with enhanced capacity to engage with local government 	These indicators are weak and can be challenged quite easily. The indicators can be made stronger by clearly outlining the areas of capacity support, as in the third indicator.
Direct measurement	<ul style="list-style-type: none"> # of successful funding proposals produced each year # of partner organisations integrating gender equality into their programmes # of NGOs establishing formal relationships with government bodies 	If support is provided to a portfolio of organisations on the same subject (in these cases producing funding proposals, integrating gender equality or establishing formal relationships with government) then developing indicators should be easy.
Action plans based on organisational assessments	<ul style="list-style-type: none"> # and description of organisations showing enhanced capacity in one or more areas of their action plans # of organisations pursuing a capacity development action plan at least one year after the start 	Action plans for individual organisations may all be very different so there may be little consistency in the indicators. As a result, portfolio indicators may need to be very broad, or may need to be based on pursuance of the plans themselves.
Organisational Assessment tools / scorecards	<ul style="list-style-type: none"> average capacity score against areas of M&E, human resources, leadership, etc. # of organisations showing an increase in capacity score in at least one area of support 	By their nature, organisational assessment (OA) tools are particularly conducive to the development of quantitative indicators. The key is more to enable a level of consistency of information collection that will make such indicators useful
Outcome mapping	<ul style="list-style-type: none"> % of organisations where at least 60% of expect to see markers and 30% of like to see markers are realised % of outcomes (represented by expect to see, like to see and love to see markers) realised 	These examples can be used across a portfolio even if every outcome map is individually tailored to different organisations. If the outcome map itself is consistent than specific indicators such as ' <i>% of organisations that develop gender policies</i> ' can be used
MSC / Case studies based on approved sampling methodologies	<ul style="list-style-type: none"> # and description of cases where organisations have enhanced capacity to engage with their constituents # and description of cases where organisations can demonstrate cultural change 	If the methodology for producing stories is transparent and valid, and stories are properly generated, then general indicators such as this can be used and justified. If MSC is used then a consistent domain can aid summarisation
Surveys / client satisfaction forms	<ul style="list-style-type: none"> # of agencies that have been asked to contribute to public fora in the past year Number of trained practitioners applying new skills effectively post-training # of supported organisations that are 'very satisfied' or 'satisfied' with support 	If surveys are applied consistently then almost any question can generate specific indicators. But to be valid the response rate for the surveys must be reasonable high, and not significantly biased.

Looking behind the numbers: While the examples shown above (many of which are taken from real-life examples) are useful for reporting to donors, they will often reveal an incomplete or misleading picture on their own. Rather than telling the full story, indicators like these tend to show broad trends and anomalies that can then be followed up in-depth.

For example, an organisational assessment could highlight areas of interest that could then be explored in depth – such as areas where a great deal of capacity change is being reported or areas where things seem to be getting worse. Likewise, a set of outcome maps might reveal that very few organisations are seeing ‘expect to see’ changes but many are seeing change at a higher level; and this may be worth exploring more intensively with more resources.

In fact many cases have been documented where perceived increases in capacity have led to lower capacity ratings (because of enhanced awareness of an organisation’s limitations) and there are also cases where capacity has increased but with no verifiable contribution from a supporting agency. Numbers will tell one side of the story, but it is almost always desirable to perform more in-depth and focused qualitative assessment at targeted points to dig for deeper and more meaningful findings.

Assessing the cumulative effect: When supporting individual organisations with capacity building it is sometimes possible to assess what effect any changes in capacity are having on a target population. The same could, in theory, apply to a portfolio of work.

If a capacity building provider is providing support to a range of organisations working in the same sector or location then it might be possible to show how changes in the capacity of a range of different organisations are having a cumulative wider effect, such as changing perceptions of CSOs in a locality, enhancing livelihoods of beneficiaries, or contributing to changes in civil society space. For example, CSSP in Ethiopia believes that its work encouraging a range of CSOs to proactively engage with local governments has had a measurable effect on relations between government and CSOs at local level.

Of course, this is not always true, and many capacity building providers provide support to organisations that are also receiving support from many other providers – or alternatively they provide support to a range of organisations working in different localities and sectors where change cannot easily be summarised. In such cases the correct approach would be to carry out larger, multi-agency studies to assess cumulative changes resulting from different agencies’ capacity development support. This would then leave each individual agency to justify its own particular contribution. Sadly, too few of these kinds of studies have been carried out to-date.

CONCLUSIONS: INTRAC’s Top Ten Tips

INTRAC has been engaged in, and writing about, the M&E of capacity development for over twenty years now. During that time the development industry has changed a great deal. Yet the basic principles for monitoring and evaluating capacity building have not fundamentally changed in that time. Based on INTRAC’s cumulative experience over the years, the following would be our top tips.

- Be clear about the purpose of capacity building. Capacity building providers need to have a clear rationale for their work, and a clear idea of what they want to achieve. This might mean developing an appropriate theory of change. At the least it should involve producing agreed statements about how improved capacity at different levels should contribute to wider development goals.
- Be clear about the purpose of M&E. M&E designed for accountability to donors and supporters is not necessarily the same as M&E designed to learn and improve.
- Decide how far you intend to measure change. For some forms of technical capacity building it should be possible to measure wider changes resulting from capacity change. For more programmatic or general capacity building it may be enough only to illustrate wider changes with a few stories.
- Carry out M&E alongside capacity building support wherever possible, and ensure that any M&E processes support the capacity building process itself (or at the very least do no harm).
- If a donor or donors are involved, agree key issues beforehand wherever possible. This should include agreeing how far M&E should go in terms of measurement, and at what levels. It might also involve agreeing what you as a capacity building should be accountable for, and what lies beyond your control.
- Alongside supported organisations, select a blend of tools, methodologies and approaches that will help provide a picture of what is changing (or not) and why.
- If you are working with a portfolio of organisations, try and develop a consistent approach to M&E that will allow you to monitor and evaluate outcomes (change) as well as outputs. Recognise that sometimes this may mean imposing specific tools or approaches on recipient organisations, and try and get their agreement and cooperation as far as possible.
- If you need to develop indicators at a portfolio level then make sure these are closely linked with the M&E tools and approaches you wish to pursue. It is usually better to identify the tools and approaches first before developing the indicators.

- Fight the battles that are worth fighting. In the current climate it is unlikely that any capacity building provider that supports multiple organisations or individuals will be able to get away with purely qualitative or anecdotal reporting. It is generally easier to develop numbers from qualitative information than to spend vast amounts of time and effort trying to persuade a donor that it cannot be done.
- Don't promise what you can't deliver. M&E staff are put under serious strain where capacity building providers attempt to prove they have achieved unrealistic expectations spelled out in logical frameworks or project proposals in order to access funding. In particular, capacity building providers should be cautious about predicting the pace of change within organisations they may influence, but over which they have no absolute control.

Further reading and resources

A wide selection of Organisational Assessment Tools are described in annex two of Simister and Smith's paper on Evaluating Capacity Building (see Simister and Smith 2010). Readers interested in MSC and how it can be applied to change at organisational level should access the Davies and Dart (2005) guide outlined below. The most comprehensive guide to Outcome Mapping is a guide written by Earl et. al. in 2001 (see reference below). This is available at http://www.outcomemapping.ca/download.php?file=/resource/files/OM_English_final.pdf.

References

- Barefoot Collective (2009). *The Barefoot Guide to Working with Organisations and Social Change*.
- Davies, R. And Dart, J. (2005). *The 'Most Significant Change' (MSC) Technique: A guide to its use*. April 2005.
- Earl, S., Carden, F. and Smutylo, T. (2001). *Outcome Mapping: Building learning and reflection into development programmes*. Evaluation Unit, International Development Research Centre (IDRC), Ottawa, Canada.
- Engel, P., Keijzer, N and Land, T. (2007). *A Balanced Approach to Monitoring and Evaluating Capacity and Performance: A proposal for a framework*. Discussion paper no. 58E, ECDPM.
- James, R (2009). *Just Do It: Dealing with the dilemmas in monitoring and evaluating capacity building*. INTRAC, UK.
- Lipson, B. and Hunt, M. (2008). *Capacity Building Framework: A values-based programming guide*. INTRAC, UK.
- MacDonald, N (2015). *Review of outcome mapping in the Climate and Development Knowledge Network*. CDKN.
- Simister, N and Smith, R (2010). *Monitoring and Evaluating Capacity Building: is it really that difficult?* Praxis paper 23. INTRAC, UK.
- Wrigley, R. (2006). *Learning from Capacity Building Practice: Adapting the most significant change (MSC) approach to evaluate capacity building provision by CABUNO in Malawi*. INTRAC, UK.