Praxis Paper 23

Monitoring and Evaluating Capacity Building:
Is it really that difficult?

by Nigel Simister with Rachel Smith

INTRAC
International NGO Training and Research Centre

January 2010

Contents

Few doubt the importance of capacity building in the modern era, and few would deny that effective monitoring and evaluation (M&E) is needed to support this work. Nevertheless, the monitoring and evaluation of capacity building is as much a challenge now as it was two decades ago. This paper examines both theory and current practice, and discusses some of the key barriers to progress.

The paper is primarily concerned with capacity building within civil society organisations (CSOs), although many of the lessons apply equally to organisations in the commercial or state sectors. It is based on a literature review and interviews with a range of capacity building providers based in the North and South. The research did not include interviews with organisations that are primarily recipients of capacity building support.[1]

The paper begins by looking at some key concepts in both capacity building and M&E. It examines different ways of thinking about M&E, and describes a variety of different tools and approaches used to plan, monitor and evaluate capacity building work. It goes on to discuss M&E in relation to donors and provides an outline of current practice, based on the interviews. Finally, it highlights key areas for further discussion, and presents some conclusions based on the research.

The main findings of the research are that where organisations are clear about what they want to achieve through improved capacity (or capacity building) and where there is a clear understanding of the purpose of M&E, it is not difficult to come up with a sensible blend of tools, methodologies and approaches that can meet the needs of different stakeholders. But if capacity building providers lack an adequate theory of change; if they do not know what results they want to achieve; or if M&E work is burdened by uncertain, conflicting or unrealistic demands, then the whole area can appear to be a minefield.

The paper concludes by presenting some practical guidelines that might help those wishing to develop or improve M&E processes, whether for learning or accountability purposes. It also highlights the importance of internal commitment to M&E at senior levels within capacity building providers. Finally, it asks whether we need to improve the incentives for those organisations that seriously wish to move the debate forwards.

1. Key concepts in capacity building

Good M&E is dependent on good planning. If the monitoring and evaluation of capacity building is to be effective it is important to know what the purpose of capacity building is, who the providers and recipients of capacity building are, and whose perspectives we are interested in. Only then can the various M&E alternatives be considered.

Basic definitions

One of the key challenges for anyone involved in the M&E of capacity building is to agree what is meant by the term. This is not easy, as there are many different definitions, some of which are contradictory. At its most basic **capacity** can be understood as 'the ability of people, organisations and society as a whole to manage their affairs successfully' (OECD 2006, p8). **Organisational capacity** can be defined as 'the capability of an organisation to achieve effectively what it sets out to do' (Fowler et al 1995, p4).
The capacity of an individual, an organisation or a society is not static. It changes over time, and is subject to both internal and external influences. Many of these changes are unplanned. For example an organisation can lose capacity if key individuals leave or change positions within that organisation. However, **capacity development** can be seen as a more deliberate process whereby people, organisations or society as a whole create, strengthen and maintain capacity over time.

INTRAC believes that capacity development is an internal process that involves the main actor(s) taking primary responsibility for change processes; it is a complex human process based on values, emotions and beliefs; it involves changes in relationships between different actors and involves shifts in power and identity; and it is both uncertain and, to a degree, unpredictable (see James and Hailey 2007).

If capacity development is understood as an internal process, **capacity building** is more often understood as a purposeful, external intervention to strengthen capacity over time. However, despite its ongoing commitment to capacity building, the development community is not clear what is meant by the concept,

---

[1] A list of the organisations and individuals interviewed can be found in annex 1.

and different organisations have different interpretations. This can lead to misunderstandings and confusion. For the sake of clarity within this paper it is assumed that capacity building involves some kind of external intervention or support with the intention of facilitating or catalysing change. The focus of M&E is therefore not only capacity development (changes in capacity at individual, organisation or societal level) but also the extent to which this is supported (or hindered) by external interventions.

A range of different players provide capacity building services. These include donors, international NGOs (INGOs), southern NGOs, specialist capacity building service providers based in the North and the South, academic institutions and individual organisational development (OD) advisers and facilitators. These providers do not always act in isolation. For example, a donor might provide money to an INGO based on its perceived ability to add-value through capacity building or other forms of partnership. The INGO might then advise a supported partner based in the South to seek assistance from a sister NGO, or it might commission an OD consultant to do capacity building on its behalf.

There is also a range of different capacity building recipients[2]. This includes individuals, organisations, and sector, thematic, geographic or issue-based networks and coalitions. Increasingly, institutional donors are also supporting capacity building at government and civil society levels; not only to improve performance directly but also to increase accountability and mutual engagement in policy making under a governance agenda. One of the first challenges for anyone wishing to design effective processes to monitor and evaluate capacity building is therefore to establish whose capacity is the focus of that M&E, and where the external support comes from.

Different perspectives

It is important to distinguish between **inside-out** and **outside-in** perspectives of capacity development. The inside-out perspective suggests that capacity development depends on an organisation's ability to effectively define and achieve its own goals and objectives (or accomplish its mission). This suggests that M&E needs to be based around self-assessment and learning in order to improve future performance, and that the organisation concerned is in the best position to know what its capacity is, what capacity it lacks, and what changes are required to bridge any perceived gaps. Outsiders may have a role in supporting this process, but any ultimate judgement on change, and the relevance of that change, must come from within.

The outside-in perspective is quite different. This suggests instead that the capacity of an organisation is the measure of that organisation's ability to satisfy its key stakeholders. In other words, the best judgement of capacity must come from the outside. This implies that self-assessment alone is not enough and that there needs to be some critical, external assessment. However, although the outside-in perspective might suggest that an organisation's beneficiaries should provide external assessment, in reality it is often those with the power and money whose voices are heard the loudest.

Another important issue is whether capacity building is supply or demand driven. If an organisation develops its own capacity building programme to address its own needs the capacity building can be seen as **demand driven**. In reality, however, the driver for change often comes from the outside – frequently from donors or international NGOs. The capacity building is then perceived as being **supply driven**.

> **Comment:** One capacity building provider based in the South, contacted as part of this research, argued that more often than not organisational assessments are carried out at the request of the donor. This can lead to limited commitment on behalf of the organisation concerned. On the other hand, they argued that when an organisation itself recognises the need to change or conduct an internal assessment the outcome is usually far more successful, and changes are often realised even where there is limited money available.

Supply driven capacity building can come in more subtle forms. Many INGOs implement programmes through local partners. In some circumstances, a certain amount of capacity building is included as part of the package. Recipient organisations know that a consequence of accepting funding is that they must

---

[2] This paper has used the term 'recipients' as a generic term for individuals, organisations or networks that receive capacity building support. The term is used for the sake of convenience and does not imply passive receipt of support. In different contexts, people may prefer to use terms such as users, clients, partners or beneficiaries.

agree to a certain level of capacity building support, and are often quite happy to receive that support. However, CSOs based in the South may act as an implementing partner for more than one INGO or donor, and therefore receive capacity support from a number of different directions. As well as the potential confusion resulting, this makes M&E much more complicated as there may be a number of different capacity building providers, all with different motivations, methods and ways of working.

Capacity building for what?

At the organisational level, capacity building is carried out for a variety of different purposes. Broadly, these can be divided into two. **Technical capacity building** is aimed at addressing a specific issue concerning an organisation's activities. Technical capacity building would not normally be expected to involve an organisation in a fundamental process of change, and would be unlikely to touch on the culture, vision, values or other core elements of that organisation. Technical capacity building is often carried out in the context of a specific project or programme with which an organisation is involved.

> "More and more when northern NGOs start a project with their southern partners, a capacity-development effort will be integrated in the activities. This means that the relationship between the northern and the southern partner is basically a one-to-one relation, meaning that the capacity-development efforts will be specific for each partner in the project, even if there are multiple partners. The capacities of each single partner are analysed and, based on this analysis, measures are taken to improve the existing capacities." (Stevens undated, p24)

General capacity building, on the other hand, is provided to help organisations develop their own capacity to better fulfil their core functions, and achieve their own mission. This type of capacity development can be slow, complex and continuous, and can require in-depth reflection on an organisation's culture, values and vision. The ultimate goal of such work is to improve the organisation's overall performance and its ability to adapt itself within a changing context. This type of capacity development is not limited to immediate practical needs (ibid).

The difference between the two types of capacity building is sometimes described as the difference between capacity building as a **means to an end** and capacity building as an **end in itself**. The table below shows that capacity building can have a range of different purposes, depending on the context. It is important for M&E that these purposes are clear, as otherwise it can be difficult to design appropriate M&E approaches. This implies the need for capacity building providers to have adequate theories of change that set out both how organisation(s) change and what the results of those changes might be.

	Capacity building as means	Capacity building as process	Capacity building as ends
Capacity building in the NGO	Strengthen organisation to perform specified activities	Process of reflection, leadership, inspiration, adaptation and search for greater coherence between NGO mission, structure and activities	Strengthen NGO to survive and fulfil its mission as defined by the organisation
Capacity building in civil society	Strengthen capacity of primary stakeholders to implement defined activities	Fostering communication: processes of debate, relationship building, conflict resolution and improved ability of society to deal with difference	Strengthen capacity of primary stakeholders to participate in political and socio-economic arena according to objectives defined by them

Source: Eade (1997, p35)

A **theory of change** at the organisational level might cover the different ways in which organisations change (see example below).

> **Example**: Reeler (2007) describes three different kinds of change and argues that the type of change considered has profound implications for M&E.

- **Emergent change** describes the day-to-day changes that are brought about by individuals, organisations and societies adjusting to changing circumstances, trying to improve what they know and do, building on what is already there, and constantly learning and adapting.

- **Transformative change** occurs when an organisation becomes stuck or goes through a period of crisis, either through natural processes or external shocks. In this case the change process is one of unlearning inappropriate ideas and values and adopting new ones in order to create a new situation.

- **Projectable change** is the kind of change that can be planned in advance, and made the focus of a specific project or piece of work. It is more about working to a plan to build on or negate visible challenges, needs or possibilities.

On the other hand, different theories of change can also be used to describe how organisational change contributes to wider aims and objectives. Ortiz and Taylor (2008) stress the importance of organisations having a clear understanding of how change happens. They argue that this means understanding the demands or needs of primary stakeholders, and the conditions required to support the emergence of change, as well as understanding the broader socio-economic environment. Put simply, if capacity building is being done then organisations need to know why it is being done, what it involves, how change is expected to occur, and how changes at individual or organisational level might contribute to any desired wider changes. An example of a simple theory of change can be found in the table below.

Comment: 'VBNK's holistic approach to capacity development is based on a set of assumptions that underpin our theory of change: when we provide quality learning services we enhance the ability of individuals to promote learning (their own and others). This in turn will lead (i) to more transparent and accountable management of development organisations; and (ii) to improved effectiveness and quality of development practice and services in the social development sector. These two outcomes are precursors to the ability of the social development sector to more effectively contribute to positive social change.' (VBNK 2009, p5)

Theories of change do not need to be very complex, and indeed from the M&E point of view they should not be. However, in their review of development literature on the M&E of capacity building, Ortiz and Taylor (2008, p24) point out a dilemma:

> "Many development organisations consider [capacity building] a fundamental part of what they do, yet very few understand what it is in a strategic and operational manner. They sense intuitively what it is. They know they do [capacity building] and why it is important (and spend large sums of money on doing so) yet they rarely conceive of it, operationalise it, or measure it in a way that helps them learn and improve their approach."

Good M&E is dependent on good planning. In turn, good planning may depend on a clear vision of what an organisation is trying to achieve. If organisations lack adequate theories outlining why capacity building is being carried out, and what the eventual results might be in terms of both organisational and societal change, it is not surprising that so many struggle to effectively monitor and evaluate capacity development and capacity building work.

2. Key concepts in M&E of capacity building

This section looks at some broad concepts around the M&E of capacity building. It examines the purpose of M&E and discusses both challenges and criteria for good practice. It also discusses how far down the chain of results (or along the ripples) M&E should attempt to measure change. Finally it looks at different directions for M&E.

M&E for what?

If organisations are to carry out effective M&E around capacity building, a key first question to address is "what is the purpose of that M&E?". The usual answer to this is a combination of accountability and learning in order to improve performance. But it is not always that simple. This is for two main reasons:

- M&E carried out to learn and improve performance will not necessarily meet the needs of accountability, and vice versa. There may be significant differences in the type of information collected, the methods used to collect it, and the honesty and integrity with which information and analyses are presented.
- There are likely to be competing demands on M&E within and across different organisations. For example, a donor might need information on the short-term results of capacity building efforts in order to be accountable to Parliament or the public. A capacity building provider might want to report results to donors, but may also want to learn in order to improve its services. The recipient of capacity building may be more interested in monitoring and evaluating their own capacity for learning purposes. And programme/project officers within that recipient organisation might simply need information for basic programme management.

The challenge is often to reconcile all these competing demands. In many cases this can best be done by ensuring that M&E meets the needs of the primary stakeholders – the providers and recipients of capacity building. Additional processes can then be introduced as required to meet the needs of other stakeholders. However, this is easier said than done, and there are often real tensions between different interested stakeholders.

It is important to note the difference between **M&E of capacity** and **M&E of capacity building**. The former is concerned with assessing the changing capacity of an organisation (or individual, or society) whilst the latter is concerned both with the quality and relevance of capacity building efforts, and the immediate changes occurring. In both cases, M&E might also be used to further look at wider changes resulting from any improved capacity.

Good and bad practice in M&E

A great deal is already known (if not always applied) about the criteria necessary for effective M&E. Some criteria are generally applicable across all M&E work. However, some are specific to the M&E of capacity building and capacity development. Some of these are described below.

- M&E is more effective for learning when delinked from funding decisions. If people feel funding or their jobs are threatened they will be less likely to provide honest and open opinions about capacity, and any changes resulting from specific interventions.
- Because the central purpose of capacity building is to enhance the capacity of those involved, it is important that M&E contributes to this process and does not undermine it (Bakewell et al. 2003).
- M&E needs to be pragmatic, and the costs should not outweigh the benefits. The danger otherwise is that large, formalised M&E systems may interfere with, or undermine, capacity development itself (Watson 2006).
- M&E should be light and should not put unnecessary burdens on organisations. However, it is important to distinguish between different stakeholders. An M&E system can be light at the point of use (e.g. for an organisation wishing to improve its capacity), whilst still being significant in terms of those providing capacity building support.

It is also important to recognise some of the very real challenges associated with the M&E of capacity building.

- The duration between capacity building interventions and desired end results can be very long. For example, one Southern capacity building provider interviewed as part of the research are only now seeing the fruits of work carried out fifteen years ago. This contrasts with the expectations of many result-based management approaches that stress short-term results (ibid).
- Results may be stretched across many different organisations. There are practical difficulties in coordinating M&E work across different organisations. These may include donors, providers, recipients and ultimately intended beneficiaries.
- Capacity is not a linear process, and organisations' capacities are constantly fluctuating. Organisations (or individuals) evolve and change over time, and are heavily influenced by changes in the external environment. Change will happen anyway, so is often difficult to attribute to specific interventions.
- It can be hard to define what a positive change is. Reeler (2007) makes the point that not all changes perceived as negative are so in reality. An organisation may go through a period of crisis,

but it may be a necessary crisis that will help that organisation evolve into a stronger organisation in the long-term. Equally, an organisation may appear to some to be in a position of stability, whilst to others it appears to be going through a process of stagnation.

There are many examples of organisations that have overcome these challenges, and developed effective M&E approaches for a range of different purposes. However, it is important to recognise these challenges at an early stage so that solutions can be incorporated into M&E design.

Deciding how far to measure

One key decision is how far to go with M&E. For example, is it enough for a capacity building provider to show that its efforts have helped an organisation (or individual) improve capacity, or should providers go further and measure the wider effects of these changes? To some extent, this depends on the purpose of the capacity building support. But it also depends on what is meant by measuring change. There is an important distinction here. Some state that M&E is primarily about measurement. However, others believe measurement is too strong a word in many cases, and prefer to use words such as assess or illustrate. For example, some organisations attempt to **measure** capacity through the use of organisational assessment (OA) tools. However, because organisations touch so many lives we can only ever **illustrate** the changes that occur as a result of improved capacity.
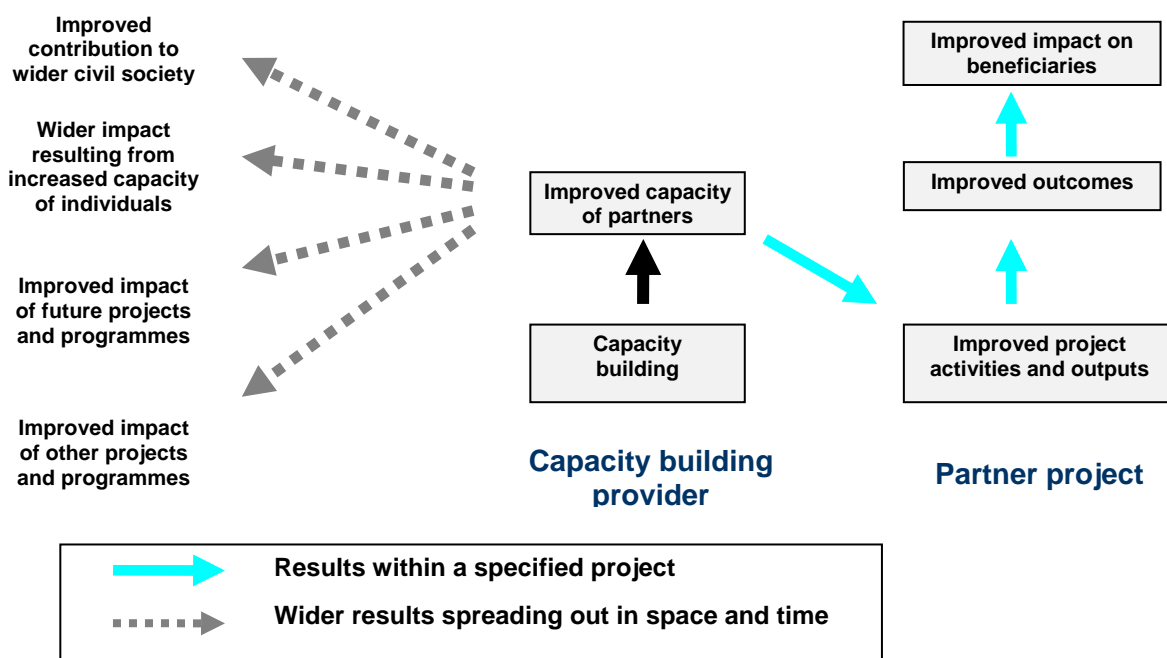
In the example above, a capacity building provider may carry out activities (such as training or mentoring) in order to support the capacity development of a partner. If this is designed to improve results in a specific project then it may be theoretically possible (albeit extremely difficult) to measure the results in terms of improved outcomes/impact at beneficiary level within that project. However, it is unlikely that benefits will be completely confined to one identified project. For example, the improved capacity may help performance in other projects or programmes run by the partner. Or individuals may leave an organisation and apply their new learning in different contexts.

If the capacity building is of a more general nature, seeking improvements in the invisible core areas of vision, values and culture, or if it is concerned with internal organisational systems such as planning, fundraising or human resources, then it will be impossible to trace all the wider results (whether positive or negative) as they spread out in time and space. In these circumstances, the best that can be done is to record some of the changes that have occurred. In other words to *illustrate* change by highlighting specific examples.

Both measurement and illustration can be effective for learning purposes. Illustrating change does not mean relying on anecdotal evidence. For example a long-term change resulting from improved capacity could be thoroughly analysed using appropriate research methodologies. This analysis might contribute
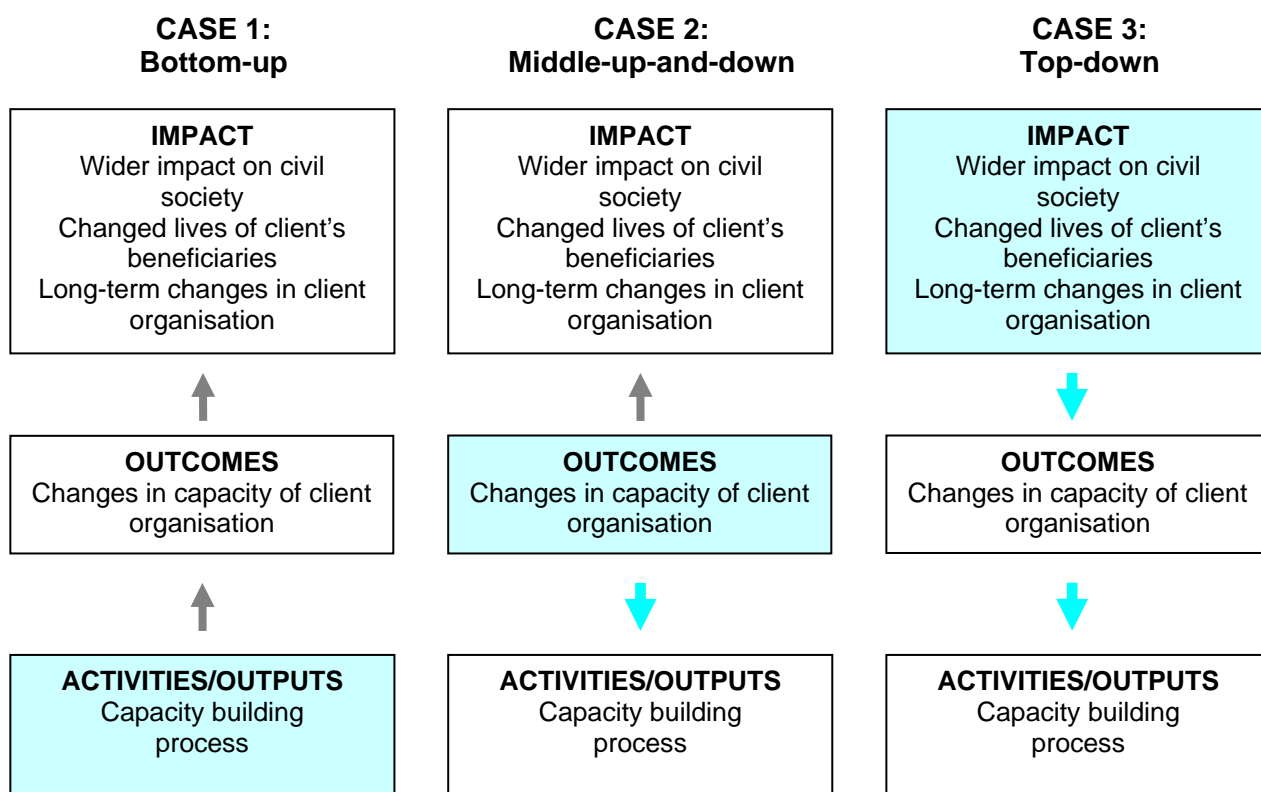
significantly to learning and improved practice. However, the recorded change will remain an illustration of wider changes. It might show a minimum change (i.e. "we have achieved at least this much") but it will not enable an organisation to comprehensively measure the wider results of any improved capacity.

Ultimately, different stakeholders need to come together to decide how far results should be measured, and where and when it is appropriate to seek illustrations of change. Agreement may be harder to reach where there is a donor to consider, but it needs to happen nonetheless. Little will be gained (and much potentially lost) if organisations pay lip service to the measurement of results in areas where it is technically and conceptually impossible.

Finally, with all the emphasis on short- and long-term results it is 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 might involve regularly reviewing and analysing the extent to which capacity building efforts are empowering or inclusive. At the very least it should involve enabling the recipients of capacity building support to say how well (or badly) they think that support was provided.

The direction of M&E

The ripple model is often used to highlight the different changes brought about through capacity building work. It shows how capacity building contributes both to changes at individual or organisational level and then wider changes in beneficiary lives or civil society. The analogy is of a stone (the capacity building input) thrown into a pond causing ripples to spread outwards. The size and direction of the ripple is influenced by, and influences, the context in which it moves (James 2009). The model is often used to show that M&E needs to focus on different levels (or ripples). But this raises the question of how to link together M&E at all the different levels. Key here is an appreciation of the direction to take. In other words, where should you start doing M&E?

CASE 1: Bottom-up	CASE 2: Middle-up-and-down	CASE 3: Top-down
IMPACT Wider impact on civil society. Changed lives of client's beneficiaries. Long-term changes in client organisation	**IMPACT** Wider impact on civil society. Changed lives of client's beneficiaries. Long-term changes in client organisation	**IMPACT** Wider impact on civil society. Changed lives of client's beneficiaries. Long-term changes in client organisation
↑	↑	↓
OUTCOMES Changes in capacity of client organisation	**OUTCOMES** Changes in capacity of client organisation	**OUTCOMES** Changes in capacity of client organisation
↑	↓	↓
ACTIVITIES/OUTPUTS Capacity building process	**ACTIVITIES/OUTPUTS** Capacity building process	**ACTIVITIES/OUTPUTS** Capacity building process

a) **The bottom-up method**. This involves starting from the support provided, and attempting to trace the changes forward. It is like starting from the pebble thrown into a pond and tracking the ripples as they spread outwards. Over a period of time M&E is used to assess:

- what capacity support was provided and to whom?
- how well was it organised and carried out?
- how was it initially received?

- what changes can be seen in the way individuals behave (if relevant)?
- what changes have there been at organisational level?
- what are (or might be) the ultimate effects of these changes on the organisation or wider population?
- what has been learnt along the way that might be of use when carrying out future capacity building work?

The bottom-up method can be used with either predictive or non-predictive M&E approaches, or a combination of both methods. **Predictive** approaches follow a logical framework approach. Goals, objectives, outputs and inputs are defined at the start of the work, and indicators are used to predict desired changes at each level. This is the most common approach for projects or programmes involving technical capacity building. In **non-predictive** approaches work is carried out and the resulting changes traced forward without relying on predicted change. This is more likely when capacity building is seen as a process or an end in itself, rather than as a means to an end.

The bottom-up method has significant advantages. Firstly, attribution is easier to assess, because M&E is focused on the results arising from a specific capacity building intervention or combination of interventions. Therefore, this is often the preferred method for donor supported work. Secondly, the bottom-up method helps ensure that the quality of the capacity building itself is included within M&E. However, the bottom-up method is less useful for evaluating the cumulative effects of different types of interventions spread over time. For example, if an organisation receives capacity support from a number of different stakeholders in the same area of its work, the bottom-up method is less suited to dealing with the complexity. Additionally, the bottom-up method makes no attempt to measure the overall capacity of an organisation. It is only interested in those areas of capacity that are being supported through capacity building.

b) **The middle-up-and-down method**. This involves making a genuine attempt to measure the capacity of an organisation at different points in time in order to show change. This is often done through the application of an organisational assessment tool (discussed in the next section). Once changes in the capacity of an organisation are identified, M&E can then be used to look backwards to investigate what might have caused these changes, and forwards to see what wider changes have been brought about.

The middle-up-and-down method may be more relevant to general capacity building than technical capacity building. It is better able to handle a variety of different capacity building inputs, applied over different timescales. For example, it would be more useful than the bottom-up method where ongoing mentoring and accompaniment are involved, as the extent of involvement might not be known at the beginning. Similarly, the method is appropriate in situations where there are many different organisations or individuals providing capacity building support to a single recipient organisation. It is also effective where there is no external capacity building support, and the only impetus for change comes from within an organisation.

One disadvantage of the method is that there is no guarantee that any particular capacity building input (such as training or a workshop) will be mentioned as a contributory factor – either positive or negative – to any organisational change. The method is therefore less useful for accounting to donors for specific capacity building inputs.

c) **The top-down method**. The third alternative is to attempt to measure change at impact level, and then work backwards to find out what might have contributed to that change. Where the enhanced capacity of an organisation (or individual) is identified as a contributory cause then it may be possible to go even further back and identify relevant capacity building inputs. This method is arguably easier to use for technical capacity building, where there is a clearly defined end-product. For example, if technical capacity support is provided to improve the capacity of traditional birth attendants, it may be possible to carry out an evaluation or impact assessment that measures changes in maternal mortality rates, and then traces this back to investigate how improved practices of TBAs might have contributed, and what might have helped bring about those improved practices.

For general capacity building the challenge is harder. Eventual impacts might include long-term changes in organisational sustainability, changes in the lives of an organisation's beneficiaries,

changes in civil society space, or changes in government or private sector policies and practices. However, organisations attempting to use the top-down method will need to have an adequate theory of change that clearly identifies what the eventual impact of capacity support might be. Otherwise it will be difficult to know where to look for long-term change. To push an analogy too far, if you want to enter a pond to measure the speed and size of ripples, and thereby draw conclusions about the initial stone that caused them, you must first make sure you are in the right pond!

The top-down method is arguably the least likely to enable meaningful conclusions to be drawn about the quality and relevance of specific capacity building inputs. There are usually a vast number of potential influences affecting any long-term change, and some warn that '[m]easuring the causes of impact within the complex processes of development can require research resources and skills far beyond the capacity of a programme's M&E activities' (Barefoot Collective 2009, p155). However, where significant M&E resources can be brought to bear, possibly through multi-agency or donor funded studies, the top-down method is arguably the most likely to show how improved capacity within different organisations can together contribute to wider changes at society or community levels.

These three methods of looking at M&E are not mutually exclusive. In an ideal world, capacity building providers could monitor and evaluate their own capacity support and attempt to trace changes forward. At the same time, recipient organisations could be supported to assess and monitor changes in their capacity and work both backwards (to see what caused those changes) and forwards (to see what wider effects they might have had). A later evaluation or impact assessment might then look at long-term changes at societal or community level and work backwards to find out what might have influenced those changes. However, where there are limited resources in terms of personnel, funding and time, organisations need to choose the approach that best suits their purpose.

3. Organisational assessment tools

Ideally, before any capacity building intervention there will be some kind of organisational assessment, whether internal or facilitated externally. An organisational assessment can be a very simple and informal exercise, perhaps involving a few straightforward questions or a SWOT (strengths, weaknesses, opportunities, threats) analysis. However, in some cases more formal tools are used to help make an organisational assessment. Organisational assessment (**OA**) tools, often known as organisational capacity assessment tools (**OCATs**) are designed to assess capacity, and plan capacity development. Sometimes they are used to monitor and evaluate capacity development or capacity building. They are the only tool in widespread use designed specifically with capacity development in mind. This section is based on the analysis of a range of different OA tools submitted by different individuals and organisations as part of this research, or identified in the literature.[3]
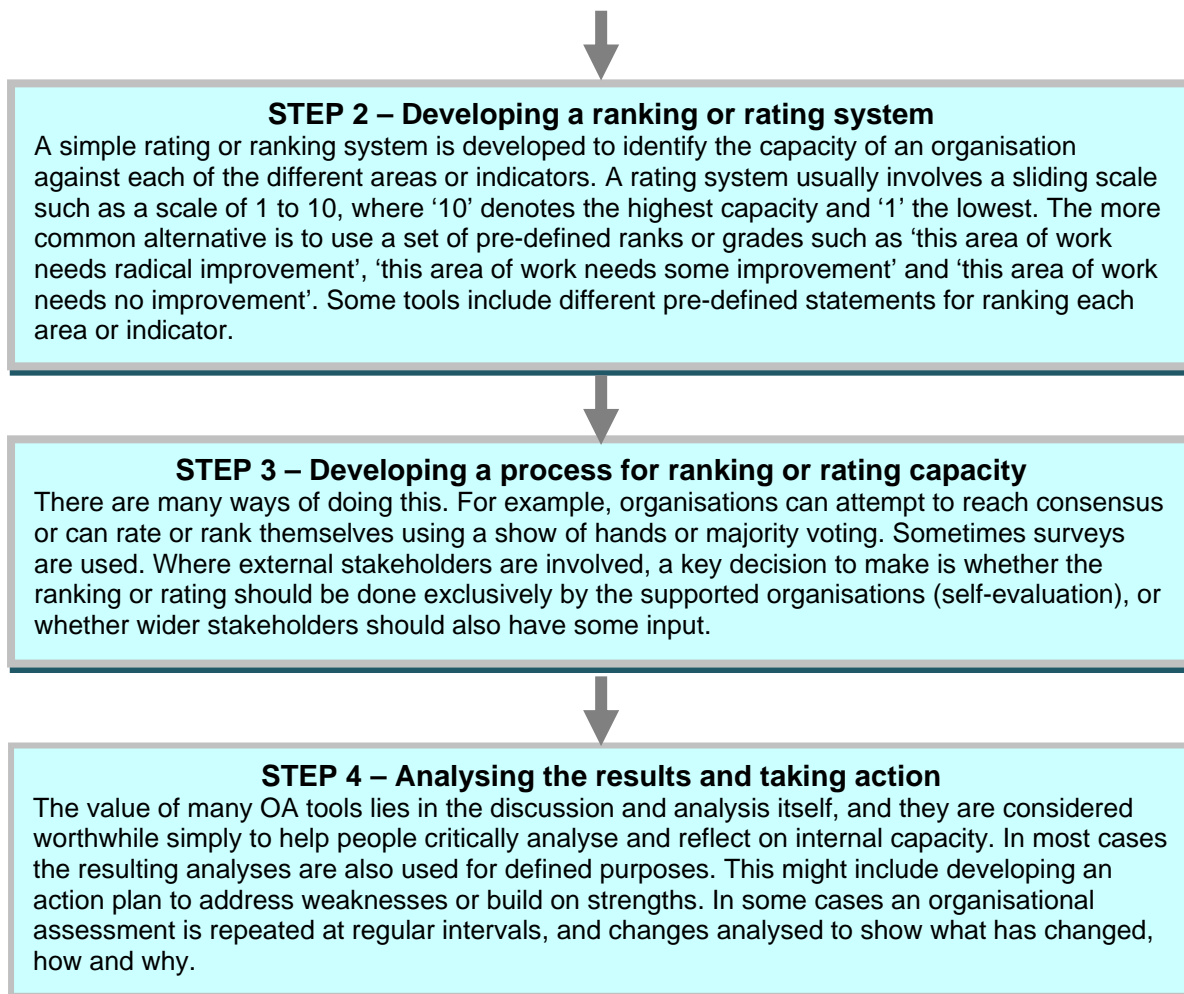
OA tools can be used in three distinct ways:

a) An OA tool may be used to assess the capacity of an organisation to act as a partner or be a recipient of funds or support. Used in this way, an OA tool performs an audit function. In these cases the OA tool often focuses on areas of capacity that are of interest to the external agency, such as financial management or project cycle management.
b) An OA tool is often used to make a general organisational assessment. It helps an organisation identify its strengths and weaknesses, and usually leads to the development of an action plan to help meet its needs.
c) Organisational assessments are sometimes repeated at discrete intervals. This is partly designed to show changes in organisational capacity over a period of time. OA tools used in this way perform a monitoring and evaluation function.

There are numerous different types of OA tools available, designed for different purposes and situations. However, most of these tools have been designed according to a similar pattern.

STEP 1 – Breaking capacity into manageable areas
Capacity is divided into a number of discrete areas. These may include areas such as internal management, relational management, ability to carry out core functions, human resources, etc. The different areas are often further broken down into more detailed statements (sometimes called indicators) each addressing a different aspect of capacity. In some tools the areas, statements or indicators are pre-set. In others there is flexibility for different areas to be defined by participants.

STEP 2 – Developing a ranking or rating system

A simple rating or ranking system is developed to identify the capacity of an organisation against each of the different areas or indicators. A rating system usually involves a sliding scale such as a scale of 1 to 10, where '10' denotes the highest capacity and '1' the lowest. The more common alternative is to use a set of pre-defined ranks or grades such as 'this area of work needs radical improvement', 'this area of work needs some improvement' and 'this area of work needs no improvement'. Some tools include different pre-defined statements for ranking each area or indicator.

STEP 3 – Developing a process for ranking or rating capacity

There are many ways of doing this. For example, organisations can attempt to reach consensus or can rate or rank themselves using a show of hands or majority voting. Sometimes surveys are used. Where external stakeholders are involved, a key decision to make is whether the ranking or rating should be done exclusively by the supported organisations (self-evaluation), or whether wider stakeholders should also have some input.

STEP 4 – Analysing the results and taking action

The value of many OA tools lies in the discussion and analysis itself, and they are considered worthwhile simply to help people critically analyse and reflect on internal capacity. In most cases the resulting analyses are also used for defined purposes. This might include developing an action plan to address weaknesses or build on strengths. In some cases an organisational assessment is repeated at regular intervals, and changes analysed to show what has changed, how and why.

Strengths and weaknesses of OA tools for M&E

In the context of monitoring and evaluating capacity building, OA tools may have a number of different strengths and weaknesses. Some of these are described below.

Strengths	Weaknesses
• OA tools can ensure that capacity development or capacity building is taken seriously, and is formally monitored and evaluated. • They enable organisations to identify necessary changes to help achieve their mission. • OA tools 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.	• 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. If there is no external input then results are open to accusations of bias. • Organisations often rate or rank themselves highly at first. Later on they might become more aware of their limitations in specific areas and might give lower scores. A lower score, therefore, does not always indicate a negative impact or failure of capacity building.

• They cover unintended or negative consequences of capacity building, as well as positive, expected ones.	• Conversely, low ranking in order to access resources may turn into higher scores at a later date with no actual change in capacity.

The value of OA tools is heavily dependent on how and why they are used. Some of the key criteria for effective use identified over the course of this research are:

- There needs to be agreement and understanding about the purpose of any organisational assessment, and how results will be used.
- If an organisational assessment is used as a tool for making funding decisions, this might encourage biased data collection and analysis and staff insecurity. For example, when VBNK, a capacity building provider based in Cambodia, used an OA tool with supported partners it made it clear that participation in the assessment was voluntary and not a condition of future funding support (Pearson 2009).
- Many OA tools work best when there is effective facilitation by an experienced facilitator.
- Many argue that unless the whole process is owned by the organisation concerned, there is a danger that the process of organisational assessment will degenerate into a lifeless technical exercise, which fails to capture reality (Barefoot Collective 2009).
- There needs to be joint analysis of findings between different stakeholders involved. Whether or not an external facilitator is involved, the value of many OA tools is derived in large part from the discussion and analysis that is involved, not from the results themselves.
- Organisations may need to have their confidentiality and anonymity respected. If assessments are based partly on individual or group interviews or questionnaires, staff members may also need to have anonymity respected.

Perhaps the biggest concern over the use of OA tools is that they are inclined to encourage a blueprint approach for organisational development. Some are critical about the practice of CSOs based 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). The fear is that emerging grassroots organisations or volunteer-based organisations are encouraged to become more 'professional' organisations, thereby losing their character as a result. In addition, standardised tools may not recognise deep contextual differences within organisations or in the wider environment in which they are based.

However, these views are not uncontested. One person interviewed felt strongly that in many countries there needs to be an agreed model of what an ideal NGO should look like, particularly where there are no established traditional roles and responsibilities. They believe that NGOs in each country should come together to decide what should be the key attributes or capacities of an NGO. This then enables self-assessment against a contextually specific, country model.[4]

> **Comment:** "Experience has shown that the exercise of deciding what an ideal NGO should look like is a very important learning exercise for the NGO, as important as the subsequent exercise of assessing the organization against the model." (PACT undated, p2).

There seems no doubt that OA tools have often been misused and abused, particularly where results of assessments have been used to deny or cut funding without fair assessment or warning. However, many organisations find them extremely useful when applied in a participatory and non-threatening manner. If we were to abandon every tool that has been misused or abused in the past we would quickly have no tools left (Simister 2000).

An evolving consensus?

Stevens (undated) argues that when trying to find indicators or statements that can apply too widely within OA tools, one often ends up with the 'largest common denominator' that can be measured in every

---

[4] Informal conversations with Richard Holloway. Any misrepresentation of his views is the result of the poor mobile signal at Gatwick airport.

organisation, but which doesn't really say anything about the organisations' capacities. The challenge therefore has been to develop capacity areas that are broad enough to apply to most organisations, yet allow for the development of sub-areas (statements or indicators) that are specific to different types of organisations at different stages of development in different sectors and countries.

One of the most generic models is the three circles model. This describes a simple model of capacity as three interlocking circles involving the internal organisation (being); external linkages (relating); and programme performance (doing) (see Lipson and Hunt 2008). More recent work by the European Centre for Development Policy Management (ECDPM) has identified five core capabilities. These, 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
- to achieve coherence.[5]

These are roughly analogous to the three circles model, with the addition of the capacity to adapt and self-renew in the future, and achieve coherence across the different capabilities. Many of those interviewed during this research have been influenced by this model, and recent work has already been carried out using the model as a lens through which organisational capacity can be assessed, and later monitored and evaluated (e.g. Phlix and Kasumba 2009).

If consensus is reached around this model (or any other model) it would go some way towards dealing with the common developmental challenge of recognising diversity in all its forms, whilst still allowing for a common framework for analysis. The model could be used to define the broad dimensions (or domains) of capacity, yet still allow organisations, or groups of organisations, to define individual statements relative to their size, status, degree of maturity and the environment in which they work. If nothing else, this would greatly simplify the task of analysing and summarising information generated through a multitude of different OA tools.

4. Other tools and approaches used for M&E of capacity building

This section describes other tools and approaches, identified during the research, used to help plan, monitor and evaluate capacity building work. None of these tools or approaches were specifically designed for capacity building work. Instead, they have been adapted by different organisations in various ways to serve their requirements.

Planning tools

The traditional method of developing a capacity building plan is to set objectives and indicators to show expected progress over a particular timeframe. This is often carried out within the context of a **logical framework** or similar planning matrix. However, many of those interviewed expressed concerns over the use of the logical framework within the context of capacity building.

- It can be difficult to develop clearly defined objectives or indicators for general capacity building work over a time-bound period, as it is often hard to predict the pace of change.

---

[5] In its final paper, ECDPM (2008) describes the five capabilities as:
1. to commit and engage: (volition, empowerment, motivation attitude, confidence)
2. to carry out technical, service delivery & logistical tasks: (core functions directed at the implementation of mandated goals)
3. to relate and attract resources & support: (manage relationships, resource mobilisation, networking, legitimacy building, protecting space)
4. to adapt and self-renew: (learning, strategising, adaptation, repositioning, managing change)
5. to balance coherence and diversity: (encourage innovation and stability, control fragmentation, manage complexity, balance capability mix)

- Capacity change can take a very long time, and most logical frameworks are designed to cover a relatively short time period.
- Any indicators defined will be dependent on the tool or methodology used to collect and analyse information. This is not always known at the start of a project or programme.

Criticism of logical frameworks is often directed at their assumption of linear, causal chains that overlook the influence of the wider environment (Ortiz and Taylor 2008). There are also concerns that, along with other results-based management systems, the logical framework tends to stress short-term changes, and does not allow enough flexibility for people to change working methods or approaches during the course of a project or programme (Garbutt and Bakewell 2005; Watson 2006).

These issues are beyond the scope of this paper. However, many of the people interviewed as part of this research believe that **outcome mapping** could be a more effective method for planning and reporting on general capacity building work. Indeed, more and more organisations are experimenting with, or showing an interest in, outcome mapping. These include donors (who are increasingly enrolling staff on outcome mapping courses), Northern and Southern capacity building service providers and OD consultants.

Many feel that outcome mapping has a number of technical advantages over the logical framework as a planning and reporting tool for general capacity building work.

- Outcome mapping requires a programme to identify boundary partners. These are individuals, groups and organisations with which a programme interacts directly to effect change, and where there are opportunities for influence. Outcome mapping is therefore particularly appropriate when assessing change at an organisational level (Earl et al. 2001).
- Outcome mapping involves the identification of a spread of possible outcomes (known as progress markers) ranging from those stakeholders expect to see to those they would like or love to see. This avoids the need for precise predictions about the pace of change at the beginning of a project or programme. However, the fact that people are encouraged to predict visible changes that may occur over a period of time means it is still a predictive tool.
- Progress markers are set separately for each boundary partner. Planning and reporting can thus be tailored individually for each separate recipient of capacity building support. This avoids the development of general indicators designed to apply across many different organisations.
- Outcome mapping focuses on behavioural change (outcomes rather than outputs). Progress markers describe observable changes in behaviours, relationships and actions of individuals or organisations that are straightforward to measure. This does not mean changes in invisible areas such as culture, vision and mission are ignored, or indeed changes in systems, physical infrastructure or resources. However, the assumption is that change in these areas will eventually translate into visible, behavioural change.
- Outcome mapping recognises complexity, and the fact that capacity building providers cannot control or force change on boundary partners, as these have ultimate responsibility for change within their own organisations (ibid).

Neither the logframe nor outcome mapping removes the need for an organisational assessment, or some other process to identify capacity development requirements. Equally, both tools need to be used in conjunction with M&E methodologies that allow for the collection (and analysis) of information defined at the planning stage.

Even diehard supporters of outcome mapping do not see it as a direct replacement for the logical framework, and indeed many organisations have successfully embedded outcome mapping progress markers into logical frameworks. The logical framework is well known, simple and convenient, and is not going to go away. For many kinds of technical capacity support the logical framework may be a more appropriate planning tool. But it is what it is – a tool designed to help plan projects, set out monitoring and evaluation requirements, and provide a brief overall summary of a project or programme. This research suggests that many people feel it is not always appropriate as a basis for planning and reporting on general capacity building or capacity development programmes. At the same time, there is increased interest in seeing if outcome mapping – whether the whole methodology is used or different elements adopted as required – could satisfy this requirement.

Stories of change

There are many circumstances where changes in capacity can be observed or measured directly. 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. However, CDRA (2001) points out that human change is often too deep and complex to measure directly. An alternative is to use **stories of change** that are capable of describing the richness and complexity of individual, organisational or societal change. Stories have long been used in development circles. However, unless an organisation is clear about how they are generated and used, such stories can be dismissed as anecdotal. In response, a number of different methodologies are used to help introduce more rigour into the process.

Along with outcome mapping, **most significant change** (MSC) is most 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 designed to ensure that the process of analysing and recording change is as participatory as possible. MSC aims to identify significant changes brought about by a development intervention, especially in those areas where changes are qualitative and therefore not susceptible to statistical treatment. It relies on people at all stages of a project or programme meeting to identify what they consider to be the most significant changes within pre-defined areas (or domains).

Most significant change was not designed specifically to support learning in capacity building programmes, but it has often been adapted for the purpose, and many users of MSC have defined domains that focus on organisational change. For example, CCDB in Bangladesh created a domain around the sustainability of people's institutions, whereas MS Denmark asked about organisational performance. Other organisations have included domains focusing on changes in communities (the Landcare support programme in Australia) or changes in partnerships (Oxfam New Zealand) (see Davies and Dart 2005).

MSC's strength lies in its ability to produce information-rich stories that can be analysed for lesson learning. MSC also involves a transparent process for the generation of stories that shows why and how each story was chosen. However, it is not designed to produce representative stories. Instead it is designed around purposive sampling – sampling to find the most interesting or revealing stories. MSC has been used by a number of different organisations contacted as part of this research.

> **Example:** 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** – randomly choosing a selection of individuals or organisations as a focus for in-depth case studies. This then allows some extrapolation of findings from qualitative information. For example, if sufficient numbers are chosen, the findings may allow for an estimation of the overall effects of a capacity building programme. However, significant resources may be required to generate enough stories to draw wider conclusions about the results.
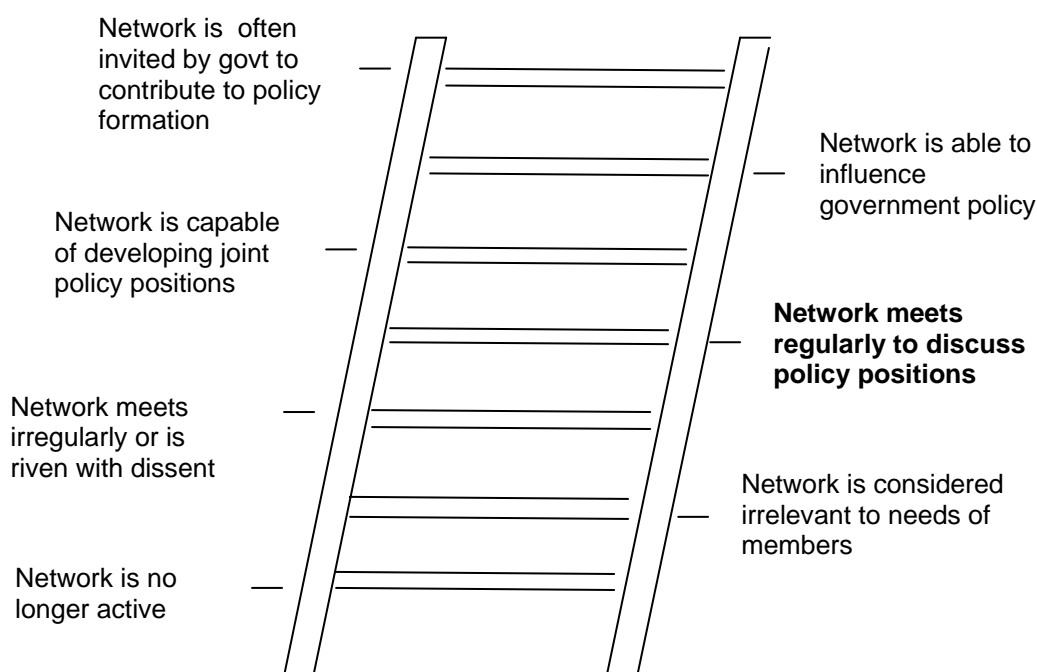
Support to individual organisations or wider society can also be assessed using purely qualitative techniques. This involves developing a qualitative baseline (a story of what the situation is now) and describing a picture of what the situation might be in the future. Regular monitoring then builds a series of pictures over time, showing what has changed and why. These are compared with the original pictures and differences analysed in order to generate learning. In essence, this is the principle of a **tracer study** – a longitudinal study providing a series of stories at discrete points in time.

Other monitoring and evaluation tools

Many other tools are used to generate information on capacity building and capacity development. These include the standard tools of M&E such as individual or group interviews, focus-group discussions,

questionnaires and surveys, direct or participatory observation and PRA techniques. Some organisations use scrapbooks or diaries to collect regular evidence of change, whilst timelines are also considered a useful method of systematically plotting observed changes or changes in opinions and impressions. Changes in individuals' knowledge and behaviour are sometimes assessed through evaluation forms, tests and KAP studies. At a wider level, appreciative inquiry is increasingly being used as a vehicle for both planning and impact assessment, and there are a number of newer tools and methodologies such as the balanced scorecard and impact pathways that are also generating increased interest. None of these tools have specifically been designed with capacity building in mind, but all have been adapted for the purpose at one time or another.

One method that has the potential to provide some rigour to the M&E of abstract concepts is a **ladder of change** (see David 1998). Ladders of change can be applied in any situation, but may be most useful when involving large numbers of organisations (for example in a network) or dealing with wider societal areas such as civil society capacity or civil society space. Developing a ladder involves sitting down with a number of different stakeholders and developing a short description of the current situation. This then becomes the middle rung of the ladder. Successive statements are then developed to show how the situation might get better or worse over time. The exercise can be repeated at regular intervals to show if change has occurred. If so, contributory factors are then investigated. A hypothetical ladder showing the capacity of a network to influence government policy is shown below (current situation in bold).

Network is often invited by govt to contribute to policy formation

Network is able to influence government policy

Network is capable of developing joint policy positions

Network meets regularly to discuss policy positions

Network meets irregularly or is riven with dissent

Network is considered irrelevant to needs of members

Network is no longer active

Some have also called for more innovative M&E techniques to be used. For example Reeler (2007, p19) argues that 'the techniques of artists, the use of intuition, metaphor and image enables not only seeing but inseeing, or the ability to have insight into the invisible nature of relationships, of culture, of identity etc.' Others argue that qualitative elements of change can be captured through participatory exercises such as drawing, characterisation and role play. However, this research did not uncover any examples of organisations widely using these kinds of alternative methods.

Client satisfaction

One of the key principles of participatory monitoring and evaluation is that whenever a service is provided one should seek the views of the intended beneficiaries. This means that the recipients of capacity building support should be encouraged to say not only whether or not their needs were met, but also whether or not the process itself was appropriate or rewarding. Many organisations have developed **client satisfaction forms** so recipients can offer a formal opinion on the value of the services provided. These include instant assessment forms (such as those used at the end of training) and periodic or end-of-project client satisfaction forms.

However, a surprising number of capacity building providers do not collect any formal feedback in this way. In these cases, M&E implicitly follows a more commercial model. The value of the services provided is

assessed by the extent to which the client comes back for more support, or the extent to which the provider's reputation leads others to seek their services. This means letting the market place define your worth. This can be seen as a valid M&E approach for demand-led capacity building work; although it could be dangerous to draw conclusions where capacity building is wholly or partly supply-driven.

Another proxy measure of client satisfaction might be the extent to which capacity building resources are accessed. Some organisations monitor how often resources are downloaded or how often websites or blogs are accessed to gauge the level of interest in their products. Where these are shown to be increasing, organisations may draw the conclusion that they are offering valuable services that meet the needs of different stakeholders.

Different M&E processes

As well as specific tools and methodologies, there are many processes widely used to share and explore different understandings of change, and to generate new findings or lessons learnt. For example, workshops, conferences and away days can be used to analyse change and generate new shared understanding of change processes. Some also see an important role for reflective reports and thought pieces that can pull together learning from capacity building work (see Pearson 2009). Research studies, internal reviews, mid-term reviews, formal evaluations and impact assessments are all vehicles through which the views of different stakeholders are brought together in order to build up a picture of change. Increasingly, INGOs are also supporting regular participatory reviews that address areas such as the impact of general capacity building programmes.

These processes can be useful in addressing the wider aspects of a capacity building programme. These include factors such as the enabling or constraining environment, relationships and power dynamics, and an analysis of different civil society actors (see Lipson and Hunt 2008). Evaluations often focus on key wider questions such as whether planning or needs analysis was appropriate, whether interventions were properly thought through, what progress, delays and insights occurred, and what would have been done differently given hindsight (see Ortiz and Taylor 2008). Evaluations or impact assessments sometimes also seek to assess the degree to which any observed changes in organisational capacity had wider impacts on targeted populations, and generate new recommendations for future capacity building (see James 2009).

Triangulating methods

ECDPM have recently carried out a large study on 'capacity change and performance' (2008). This stressed the need for many different approaches to be used in monitoring and evaluating capacity building and capacity development. The results of capacity building work can rarely be assessed through statistical methods alone, or through purely qualitative methods. Instead, there needs to be a combined approach using different M&E tools, methodologies and approaches to build up a picture over time of what has changed, why it has changed, and how learning can be applied in the future.

Some organisations combine traditional planning models, such as the logical framework, with newer methodologies such as outcome mapping or MSC. Some combine regular organisational assessments with periodic reviews or more formal donor evaluations. And many use different methodologies to gauge the opinions of a variety of different stakeholders throughout the chain of support from donors to communities. However, organisations also need to carefully assess their planning and M&E needs against the requirements of different stakeholders and the resources available to carry out M&E work. Theoretically, there are enough different tools and approaches to enable any organisation with sufficient commitment (and resources) to build up a picture of change. The challenge is 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.

5. Donors

Over the past few years donors have invested enormous amounts of money in capacity building and for many it is seen as a strategic priority. Even when capacity building providers receive income through charging for services, recipient organisations often pay either directly or indirectly with donor money. Under

these circumstances donors inevitably wield a large amount of influence over how capacity building is monitored and evaluated.

Accountability for what?

M&E is often discussed in relation to accountability. However, it is important to recognise that accountability covers a wide range of different areas, including joint objective setting, transparency of decision-making, financial accountability, open and honest dialogue and a host of other factors. Indeed it is perfectly possible for an organisation to be held accountable purely for the quality of its learning, reflection and improvement processes. Reporting on the basis of M&E is therefore just one aspect of formal accountability, albeit an important one. In this context it is helpful to look at different levels of accountability.

At the most basic level, capacity building providers can be held accountable for **activities** and **outputs** (i.e. what they do and produce). This includes accounting for money spent, spending it on what it is meant for, and trying to ensure that any work carried out is both the right thing to do and is done as well as possible.

Capacity building providers can also be held accountable through their **outcomes**. This is a more difficult area, as donor agencies need to ensure that organisations are open to innovate, take risks and work in areas where outcomes are hard to achieve (or measure). However, most people believe it is reasonable to expect capacity building providers to report on initial changes arising out of their work, whether positive or negative. This means attempting to find out, and report on, changes within organisations or individuals who are the direct recipients of capacity building work.

However, accountability through **impact** brings in a host of new problems:

- Firstly, different parties understand the term in different ways. James (2009) points out that for capacity building providers, impact is often seen as change at the organisational level of a client or partner. A donor, however, might see impact more as change at beneficiary or wider civil society levels.
- Secondly, impact on beneficiaries or wider civil society may not be seen until well after the timeframe of a typical project or programme. By the time impact occurs there may be no money to carry out M&E work, and little interest in pursuing it anyway.
- Thirdly, impact can be impossible (or at least extremely difficult) to measure. So accountability at impact level actually means accountability for measurable impact – which might unduly influence the kind of work capacity building providers are prepared to do, or the kind of organisations they are prepared to support.
- Finally, the international development community has for years been encouraging organisations, especially INGOs, to work through the development of Southern partners. This is based on the assumption that capacity building leads to sustainable benefits. Is it the fault of an INGO if the theory does not work in practice, or if the benefits to communities take longer to materialise?

At the same time, donors argue that they need to see a return on their investment. They, too, have stakeholders to which they are accountable. For institutional donors these may include politicians, parliamentary committees, national audit offices and ultimately the public. They need to demonstrate that their funding is contributing to poverty eradication or realisation of human rights. And they may need to do it in a way that can be clearly understood by people with no understanding of the complexities of international development, especially in the current economic climate where Northern government spending is under increasing scrutiny from the media and public.

Crucial here is the difference between M&E as measurement and M&E as illustration. It may be difficult or impossible to *measure* wider changes resulting from capacity building work. However, it is reasonable to expect some *illustration* of at least some of these changes. This then raises the dilemma of who should carry out M&E work at impact level. Some say it should be the capacity building provider. But they may argue they have limited access to beneficiaries, or that measuring results at beneficiary level might undermine their clients or partners. Others believe it is the responsibility of the recipient organisation, whilst many argue that the donor needs to be involved in any assessment of impact.

The heart of this debate lies in the question of where measurement stops and plausible assumption (backed up by illustrations of change) should take over. After all, if an organisation can show that it is

reducing the spread of HIV, or reducing morbidity rates, no one expects it to go further and prove the wider impact resulting every time. Even credit programmes sometimes measure no further than the disbursement of loans, on the assumption that the likely impact of such loans is already known. If, then, the development community has decided that improved capacity is likely to have a positive effect on development, it is not reasonable to expect capacity building providers to test this assumption on every single occasion.

For general capacity building at least the counter argument is that we simply don't know enough about whether or not the improved capacity of Southern-based organisations leads to improved lives, and how. Consequently, we are less free to make such assumptions. In order to acquire such evidence, some suggest a valid approach would be to undertake large, possibly multi-agency, studies to test the assumptions, and arrive at a better understanding of the links between improved organisational capacity and long-term impact. Such studies might be difficult to arrange and fund. They might also be controversial if they were to seriously test some of the assumptions surrounding capacity building. However, if done well, such studies could potentially provide rigorous evidence linking improved capacity to improved impact.

> **Example:** The Aga Khan Development Network (AKDN) is currently undertaking a baseline of what civil society looks like in eight countries. The purpose of the study is to overcome the current weak understanding of civil society and its challenges. AKDN intends to repeat the exercise at regular intervals to find out what has changed, and to analyse contributions towards those changes. The intended outcome is better understanding by government, business and the public of the breadth and value of CSOs in each country, and existing blockages to their performance. The study will involve collaboration with CSOs and other key institutions. AKDN intends to use the work to enhance capacity building efforts within each country. If successful, the study could help to show clear links between capacity building efforts and improvements in the contribution of civil society (AKDN 2009).

However, there is another donor perspective that almost negates the whole debate. This is that capacity building providers receiving donor funds should simply measure what is in the logical framework. This means measuring performance at the purpose or specific objective level of a logical framework, and no higher. Those supporting this perspective argue that a donor agency decides whether or not to fund a project or programme based on a submitted logframe. If this logframe does not include objectives relating to wider impact on beneficiaries or civil society then organisations are not expected to measure such impact, and vice versa. This perspective is indicative of a view that sees project/programme funding (and therefore M&E) as an instrumental approach to be dealt with on a case by case basis.

Quantification

There is an ongoing debate concerning the relative values of stories and numbers. People are often artificially divided into two camps – those who value stories most (whilst accepting that numbers are sometimes necessary) and those that value numbers (whilst recognising that stories are also important at times). Not all donors require extensive quantitative data. However, many capacity building providers report that they are coming under increasing pressure to justify funding by providing quantitative data at outcome or impact level. This pressure may be both internal and external. For example, in the UK there is increasing pressure from DFID on INGOs to provide quantitative measurement of change. In many cases, this pressure is mirrored by internal views at senior management level.

This research also identified a third camp. Some are increasingly frustrated by the debate and simply don't see what the fuss is about. They argue that if donors want numbers then give them numbers. This is significantly easier when organisations are providing technical capacity building for a defined purpose. However, even with general capacity building it should not be beyond any capacity building provider to quantify at least some of its results. On the contrary, whilst individual cases of capacity building support can easily be evaluated through purely qualitative measures, summarising progress across a number of recipient organisations almost inevitably involves some quantitative presentation of results, progress or lessons learned.

Numeric data can be generated through many of the tools and methods described in previous sections. For example, OA tools are inherently numeric, and use ranking or rating systems that can be analysed to produce statistics. (It is true that a common complaint is that assessments of capacity can go down over time as organisations increase understanding of their limitations; but this can easily be overcome either

through counting all changes, or by investigating each recorded change to see if it is positive, negative or neutral). Other tools and methodologies that can be used to generate numeric data include workshop or training evaluations, ladders of change, surveys, client satisfaction forms and records of people accessing capacity building resources.

Where organisations focus on stories of change there is also plenty of opportunity for quantification. If stories are based on random or representative sampling then qualitative findings can be extrapolated to generate numbers. Even techniques like MSC, which uses purposive (and therefore not representative) sampling, have clearly developed methodologies for generating quantitative data (see Davies and Dart 2005). In fact, one of the largest advocates of qualitative methodologies over the past two decades argues that there is much unrealised scope in this area.

> "Participatory methods have a largely unrecognised ability to generate numbers which can also be commensurable and treated like any other statistics. Through judgement, estimation and expressing values, people quantify the qualitative. The potential of these methods is overdue for recognition." (Chambers et al. 2009, p6).

All three camps agree that a mixture of different types of information is needed to present a full picture of change, although they might disagree on the precise balance. But where there are internal or external requirement for numbers then there seems no conceptual reason why they cannot be provided. The consensus amongst those interviewed as part of this research seems to be that if some numbers can be provided at outcome level, together with stories providing illustrations of change and where possible explaining how change occurs, it would take a fairly unreasonable donor to demand more.

However, there are two important considerations. Firstly, the value of numbers derived from qualitative methodologies depends heavily on the skills and integrity with which those methodologies are pursued. Numbers produced from poorly designed or implemented methodologies are likely to be meaningless at best and misleading at worst. Secondly, in order to generate quantitative data from qualitative data one must first have carried out work to generate the latter. There is a suspicion that those who complain the most about having to generate numbers are those that currently produce neither effective quantitative nor qualitative information.

Moving goalposts

Donors are on the move, but in which direction? Many people interviewed as part of this research – donors and recipients – stated that donor demands for (usually quantifiable) evidence of results are increasing. This is not confined to governmental donors – many of the larger INGOs are also demanding more results-based M&E. Some have concerns that this trend could inhibit learning-based approaches to M&E, which encourage feedback, lesson learning and improvement rather than measurement (see Watson 2006).

However, any attempt to acquire the 'donor view' needs to recognise that donors are not monolithic beings, and there is often a wide variety of views within any single donor. Where there is no defined organisational view, this can cause problems. More than one capacity building provider contacted during the research had experienced problems with changing demands on M&E following a change in donor personnel midway through the course of a capacity building programme or project.

Yet examples were also provided of donors that expected little formal M&E (such as some of the philanthropic donors), were happy to negotiate around M&E expectations, or were happy to accept purely qualitative reporting as a vehicle for accountability. Some people interviewed felt that there have been positive changes within the donor community over the past few years, and that it was important to tap into these changes. One example frequently provided concerns the ECDPM project, which was carried out with the support of a variety of institutional donors to develop some consensus around what works (or doesn't) in capacity building. If the conclusions of this study are widely accepted then there is the prospect in the near future of a change away from formal planning models and technocratic approaches to capacity building, and towards more experimental and incremental approaches (ECDPM 2008).

As organisations providing capacity building services are increasingly funded through a variety of different sources, it becomes more and more important to have some accepted frame of reference within which M&E can take place. Without this it can be extremely difficult for those pushing for real M&E change from

within organisations to have their voices heard. Whilst the fear of donors (or changes in donor demands) persists the views of those who see M&E largely as a vehicle for providing formal accountability and raising funds is likely to hold sway over those who see it as an important process for learning and improving.

6. Current practice

This section covers current M&E practice within capacity building providers. It is mostly based on the interviews carried out for the research. To respect confidentiality different views have not been attributed to particular agencies.

How much M&E of capacity building is carried out?

Capacity building providers can be divided roughly into two groups. The first group includes organisations that specialise in providing capacity development support. This includes capacity building organisations based in the North and South. It also includes INGOs, such as VSO, that primarily exist to raise the capacity of partners in the South. The second group includes INGOs and networks, which carry out technical and general capacity building as part of an integrated approach to development.

There is plenty of available evidence that organisations in the first group carry out significant work to conceptualise capacity building and develop M&E approaches that seek to assess change at organisational level, and sometimes wider. For the most part, attempts to assess wider change (or impact) rely on illustration rather than measurement. Planning and M&E is often based around some kind of organisational assessment, followed by monitoring of an action plan using many of the methodologies described in section 4. However, there are also examples of major capacity building providers that carry out little or no formal M&E of capacity building. The prospects for systematic and effective M&E tend to depend on:

- an appropriate theory of change that clearly spells out what an organisation is trying to achieve in the short- and long-term through capacity building support
- senior management's internal commitment to M&E for learning
- either core funding that enables resources to be devoted to M&E without passing charges onto recipient organisations, or methodologies that can be applied alongside, or as part of, the capacity building process.

Some examples of M&E approaches used by specialist capacity building providers are shown below.

Example: VBNK, based in Cambodia, has a mission to learn and improve that is internally driven. It receives core funding that allows it to pursue its own M&E approaches. Following an organisational assessment and the development of an action plan, systematic M&E starts by looking at peoples' impressions of a capacity building intervention. VBNK then looks for changes in the workplace, such as developing/applying new policies. It investigates how supported organisations deliver services to the public, and facilitates annual community conferences with the involvement of beneficiaries, NGOs, private sector and government. VBNK also carries out an annual impact assessment. The focus changes from year to year. In 2009 four key methods were used to generate information – appreciative inquiry, MSC, interviews and focus group discussions. The impact assessments are not seen in isolation, but rather as a series of reports building up an evidence-based picture of change over time. VBNK believes it has learned a lot about how to be a more effective organisational development organisation through M&E.

Example: CDRA strongly believes in a culture of internal learning and reflection. As a result it allocates funding for learning in its budgets wherever possible, justifying it on the basis of improved performance both for itself and the organisations it supports. Different capacity development practitioners are relatively free to pursue their own methodologies, and can use different M&E approaches. Many rely on continuous feedback from clients and observation of changes based on long-term contact. Individual learning is then translated into organisational learning and improved

performance through systematic procedures. CDRA staff individually produce periodic analytical reports, which are debated within the staff team, forming a kind of peer accountability mechanism. Strategic decision-making results directly from discussions and review of practice. In addition, CDRA carries out self-evaluations, and sometimes commissions external reviews of sampled work. Some programmes are also evaluated, either internally or by external consultants. CDRA believes strongly in using M&E for continuous reflection, learning and improvement rather than for reporting to external stakeholders.

Example: Pact tends to use OA tools to facilitate organisational assessments. This allows partners to assess their strengths and weaknesses along multiple dimensions of management; including strategic direction, organisational structure, governance, planning, fundraising, financial and grants management, human resource management, and monitoring and evaluation. Based on the findings, which are generally carried out with a cohort of organisations, Pact develops a tailor-made, capacity building programme, which usually combines training, mentoring, and one-on-one technical assistance. In order to measure the impact of this work, Pact generally reapplies the OA tool in the second or third year of programme implementation.

For INGOs carrying out an integrated approach to development, where capacity development is just one element of their work, capacity building needs to be divided into two areas. Where technical capacity building is carried out as part of a wider programme, M&E often relies on the development of objectives and indicators within logical frameworks or similar programme-level matrices. The extent and quality of M&E varies from programme to programme, and is often reliant on the level of detail contained within objectives and indicators. The contribution of capacity building to the overall impact of a project or programme is rarely the exclusive focus of a review or evaluation, but is often included as one aspect.

For general capacity building the picture is mixed. On the basis of our interviews, there is little evidence that most INGOs make any systematic attempt to carry out M&E of capacity building as part of a wider strategy. Some have concept documents or research documents that discuss M&E of capacity building, but these have rarely been translated into organisation-wide policies or practices. Some INGOs stated that much theorising work is now considered obsolete or is "gathering mould on shelves". In addition, there were examples of staff based in Head Offices who did not know the extent of M&E of capacity building carried out within their organisations, clearly indicating the lack of a coordinated approach. Some, indeed, regard the subject as largely passé.

The extent of interest within many INGOs or confederations surrounding the importance of capacity building often varies enormously from country to country, as does the extent of M&E. Typically, it is not dictated by central policy, but depends heavily on the focus and interests of different country offices. Examples were provided of country offices that are very interested in the subject and are actively pursuing their own approaches. Some examples were also provided of INGOs that are in the process of developing programme models which include capacity building, and are still interested in searching for new ways to monitor and evaluate capacity building.

There was also some evidence of geographic or sectoral differences. For example, more interest seems to be shown by INGO offices based in Central Europe and Scandinavia than those based in the UK and Ireland. More interest also appears to be shown in the M&E of capacity building from people working on advocacy approaches, where there is currently a wide debate around the relative merits of INGOs using their own 'voice' in advocacy work, or working slowly to help improve the advocacy capacity of indigenous organisations.

However, these examples did not appear to be indicative of wider trends. Indeed, many of the people interviewed clearly felt some level of frustration at the lack of progress made in the area of M&E of capacity building over the past few years, and thought that INGOs as a whole should be doing better.

Barriers to carrying out M&E of capacity building

The previous section suggested that work - some of it new and innovative, some based on older models – is being carried out to monitor and evaluate general capacity building. But it is patchy and inconsistent, which makes it hard to draw overall conclusions more generally across a wide range of organisations. The research is not extensive enough to draw firm conclusions, but many examples were provided to suggest why M&E of capacity building has not advanced (or is not as widespread) as it should be.

Firstly, many organisations say they lack the means to carry out M&E of capacity building work effectively. Formal M&E requires the time and effort of both providers and recipients of capacity building. It also requires money. Many organisations do not have adequate resources, or are already buried under huge reporting expectations of institutional donors. Some are not convinced that the benefits of M&E work match the level of resources required. One example was provided of an organisation that had recently developed an applied learning centre to reassess its approach to capacity building, but lack of funding meant the project has been put on hold.

Secondly, a number of organisations do not consider the M&E of capacity building to be a priority. This can be for a number of reasons:

- Many organisations' M&E systems are oriented more towards accountability (particularly to donors) than learning in order to improve performance. If these donors do not have a clear idea of desired impacts then organisations may feel they have little to gain by pushing the issue at this stage.
- For some INGOs, capacity building is just one – and not always the most important – aspect of its work with partners. Any organisation has limited resources to carry out M&E work, and some INGOs prefer to devote these resources to monitoring areas such as partnership, child and youth participation, or equity and inclusion.
- One person interviewed believes that the current pressure to focus on results - whether internally or externally driven - almost inevitably results in a loss of focus on the means. In other words, if organisations are always looking at the end-results of partners' work they lose focus on the process of how they get there. The M&E of capacity building is thus seen as far less important than the M&E of the results achieved by those partners.
- Some people interviewed discussed pilot initiatives to improve M&E of capacity building that foundered due to lack of senior management support (see example below).

Example: In one INGO a pilot programme was developed that sought to assess capacity (both individually and organisationally) to carry out advocacy work. The pilot programme was evaluated with the intention of rolling the methodology out across the organisation. Unfortunately, there was a restructuring and management support for the programme gradually faded away. The initiative was lost as a result.

Thirdly, some organisations wish to do more but feel constrained by other factors. The most common factor observed is that organisations simply don't know how to monitor and evaluate capacity building, and regard it as too difficult an area. Other factors include:

- Many INGOs have no clear rationale for general capacity building, or consistent theory of change. In particular, opinions are often sharply divided within organisations about whether capacity building should be focused on obtaining immediate results within established programmes of work or whether it should be part of longer-term efforts to improve the capacity of Southern civil society.
- Some organisations lack the staff required to carry out effective M&E work around capacity building.
- Many INGOs are currently undergoing amalgamation or restructuring, and feel it is the wrong time to be pursuing new initiatives. An example was given of a member of a confederation that has developed a new organisational assessment tool, but has been unable to implement it due to a forthcoming amalgamation.
- Some of the people interviewed desire to push the agenda of M&E of capacity building further, but do not have the power or influence within their organisations to do so.
- Finally there is an increasing tendency for INGOs to support partner capacity by putting them in touch with other partners (mentoring) or capacity building service providers based in-country. These INGOs are honest about the fact that they have neither the resources nor the capacity to undertake organisational capacity building work themselves. But if INGO staff are not actively engaged in capacity building work they are unlikely to be able to monitor and evaluate it effectively.

Some, however, argue that these are all merely symptoms of one overriding problem. So much time, effort and money has been put into capacity building that there is a genuine fear of what might be found if we look too closely. There are concerns that investments in capacity building have not brought about desired

changes, nor have they resulted in the promised impact (see James and Hailey 2007, Reeler 2007). This is an important area for debate. If M&E is not carried out because of practical concerns about resources or lack of technical know-how then it is possible to rectify the situation. On the other hand, if there is a wider malaise then it may be much harder to persuade organisations that it is in all our interests to attempt to find out the truth.

7. Questions for further debate

Based on the research, there are several areas where further debate would appear to be necessary.

What is M&E for?

If M&E of capacity building is to improve, we first need to know its purpose. There is a gap between the literature and perceived current practice. Much of the literature emphasises the importance of M&E being used to continually learn and improve (see Ortiz and Taylor 2008, James 2009, Barefoot Collective 2009, ECDPM 2008 and many others). Yet current practice in the M&E of capacity building and M&E more widely is often aimed at accountability to donors. M&E carried out for this purpose can at best inhibit the process of learning and at worst make a mockery of it. But it is hard to know what to do about it. Some have suggested that M&E for accountability and learning are never going to be compatible and there needs to be a formal separation of the two functions (Mebrahtu et al. 2007). In the absence of any clear direction from the donor community, capacity building providers and recipients will be left to make decisions on a case-by-case basis. This may not be ideal, but will at least be better than pretending that M&E can really serve two masters at once.

Standardisation of organisational assessment tools

Many have pointed out the dangers of imposing standardised, global checklists for organisations of different types, sizes and maturity, existing in different contexts and environments. Yet there are many who believe that some level of standardisation is necessary. The challenge is to reconcile the different viewpoints. The five capabilities model recently developed through the recent ECDPM project is currently exciting much interest, and could serve as a future model. This would not prevent individual organisations, sectors or countries developing their own specific indicators or statements against which to assess needs and monitor progress. However, the five capabilities model might suggest an overall framework that encourages people to think consistently about the requirements for a well-rounded organisation.

A great deal of investment has been made in the ECDPM project, and the development community now needs to decide how far to take the model forwards. At the moment it remains largely theoretical. However, as more and more organisations begin to experiment with the model in practical ways, it needs to be analysed in order to better understand its potentials and limitations. Above all, findings then need to be presented in an accessible way so that the debates are not restricted to academic circles.

The adoption of outcome mapping

There seems to be significant demand for an increase in the use of outcome mapping, either as an alternative to the logframe, or as a supplement. Even people who have little or no experience of outcome mapping are beginning to question whether it does not have a role to play in the planning and M&E of capacity building. Further practice and research may be needed in this area. However, in order to smooth the path of this research there needs to be a clearer message from the donor community about how, and in what circumstances, outcome mapping may be appropriate. At the moment there are mixed messages, as departments within some donor organisations are still insisting on rigid adherence to the logframe, whilst others are busy enrolling staff on outcome mapping courses.

In some areas the two tools may be compatible. However, the most basic difference is that the logframe asks people to predict results over a typical three to five-year period, whilst outcome mapping acknowledges that change is harder to predict, and needs to be monitored over a wider spectrum of possible changes. This tension needs more debate and more clarification, not least because many organisations still complain they lack the space to experiment with outcome mapping.

M&E of individual capacity

Training as a vehicle for capacity building has fallen off the agenda over recent years (Cracknell 2000). Consequently, less interest has been shown in the monitoring of individual capacity. Yet many international agencies complain there are insufficient local staff of suitable calibre in key areas. These include carrying out or facilitating advocacy work, providing capacity development support, facilitating community development, engaging in participatory M&E or impact assessments and a host of other areas.

Writing about the humanitarian sector, Christoplos et al. (2005, p47) warn of the dangers of exclusively focusing on organisational capacity in an environment where local staff frequently move between different organisations as better-paid, more stable or satisfying opportunities come along.

> "Paradoxically, building and investing in capacity at an individual level may be more 'sustainable' than institutional development, especially when the political and institutional context is turbulent and uncertain. The international aid community is so focused on assumptions that capacity building has to be institutional that the impact of building a strong national cadre of personnel who may move from one institution to another is overlooked."

This might imply refocusing M&E more at the level of the individual, rather than concentrating solely on assessing organisational or societal change.

M&E of wider civil society

This paper has argued that much general capacity building is aimed at promoting and enhancing civil society within different countries. But it is unrealistic and inefficient to expect every provider to carry out independent studies to assess whether or not improvements in the capacity of indigenous organisations contribute to improved civil society. Instead, there is an argument that the development community as whole should be addressing these issues. Some organisations have already started to attempt to monitor the strength of civil society in various contexts. One example is the AKDN study covered earlier in this paper. Another is the CIVICUS civil society index (CSI) - a participatory needs assessment and action planning tool for civil society around the world, designed partly to assess the state of civil society in different countries (see CIVICUS 2009).

However, we may need more such studies, properly funded and based on the involvement of a wide range of stakeholders. If such studies are able to show the link between improved capacity of Southern organisations and improvements in wider civil society – or even show some tentative links (illustrations perhaps) between broader civil society and impact on the ground, M&E of general capacity would suddenly become both more important and a lot easier to carry out. Organisations would no longer be expected to show wider impact every time they helped facilitate changes in recipient organisations' capacities. Instead they could focus M&E on progress at organisational or individual level and rely on an adequate theory, backed up by reliable evidence, of the links to improved civil society. Of course, any widespread study could show that there are no (or unproven) links. But that is the risk you take if you are serious about M&E.

Donor agreement on extent of M&E

Tied up with previous arguments, many capacity building providers simply do not know how far down the results chain they are expected to go with M&E. Most acknowledge that they need to show changes at organisational (or individual) level. But is it enough to illustrate wider changes resulting from improved capacity, or do these need to be measured? Should organisations have to show clear attribution for wider changes, or is it enough to draw plausible linkages? Should a capacity building provider have to go over the heads of a client or partner to carry out systematic M&E at beneficiary level, or should this be the job of the recipient organisation with the support of the donor?

We need to decide whether we are content to make these decisions on a case-by-case basis, depending on the purpose of capacity building, the context and the donor (or range of donors). Or whether we can come together to agree some general standards and guidelines to assist decision-making. These would do much to remove the fear factor that so often dominates decision-making when capacity building providers do not know whether expectations will change if there are changes in personnel within donor organisations.

External judgement

From the outside-in perspective there will always be some requirement for external judgement on the capacity of an organisation that goes beyond the information supplied by that organisation itself. In many cases this takes place through external reviews or evaluations that often include multiple stakeholders such as donors, capacity building providers and recipients and wider beneficiaries. However, CSOs based both in the North and South are increasingly being asked, or are volunteering, to undertake external assessments, either to generate recommendations for improvement or to acquire an external seal of approval. Some argue this is a positive thing, especially in countries or societies where there is little history of civil society development, or where CSOs are viewed with suspicion. In such circumstances external certification may involve measuring CSOs against their own criteria.

There is more concern about the tendency to assess CSOs against a long list of indicators designed to benchmark against the perfect NGO. For example, one private company is currently promoting a system that benchmarks organisations against over 100 indicators, selected from different codes and international standards. If such certification is purely voluntary then it may serve some purpose. But if organisations feel compelled to undertake such certification exercises there are dangers it will encourage the imposition of monolithic standards on Southern organisations.

Yet there is still a debate to be had concerning how far the capacity of any organisation can be judged using purely internal criteria, and whether in such cases there will always be some question over the legitimacy of the process. External evaluations, regulation, inspections, accreditation and adoption of externally developed self-regulation codes are part of the armoury for assessing change in organisational capacity, and there may be times when they are required to combat fears of subjectivity and bias. It remains to be seen whether or not these need to (or can be) held separate from M&E processes.

INGOs and added-value

For organisations specialising in capacity building, M&E is (or should be) a priority area to help them learn and improve performances. For many INGOs the position appears to be more confused, and capacity building work may be just one element of a range of different activities carried out to add value. Other elements might include promoting participation, promoting equity and inclusion, linking advocacy between different levels, networking and encouraging partnership. But there is significant overlap between M&E of capacity building and M&E of other added-value areas, and many of the methodologies and principles highlighted in this paper could be applied in these other areas as well. Indeed many of the tools INGOs have developed to monitor progress in areas such as participation and inclusion are very similar to the OA tools designed to help organisations assess their strengths and weaknesses, and plan, monitor and evaluate capacity building work. For instance, the final two examples of OA tools contained in annex 3 describe two such tools presented as part of this research.

For many INGOs, then, the whole debate around the M&E of capacity building needs to take place within the wider debate around added-value. What is it that INGOs add to the development sector? What are their priorities? Should they be carrying out capacity building work at all, or should they be encouraging more specialist in-country organisations to provide general capacity building services? Are they willing to be seen primarily as sub-contracted donors whose principal purpose is to help target institutional donor money to Southern civil society? Ultimately, scarce M&E resources need to be directed towards what is considered most important to an organisation. The sense from this research is that M&E of capacity building may be of lower priority to INGOs at present. Which begs the question: what are their priorities?

The gap between theory and practice

Finally, a great deal has been learnt over the years about the factors that enhance or inhibit good M&E of capacity building. Yet too little of this information is accessible to the practitioner. Much of the debate is couched in academic language or deals with abstract concepts and theories. There is a need to collect this information in one place, and present it in an accessible form. People seeking to develop new tools, approaches or methodologies should not, in this day and age, have to wade through the internet to find different versions of tools and papers developed or written over a twenty year period to find what they need.

The theory has moved on, but practice – as ever – is slow to catch up. To some extent this may be inevitable. But we could certainly make it easier for people to access information by holding it in one place, or summarising key lessons from the past in language that all can understand. Otherwise, M&E of capacity building risks ending up as the poor relation of other kinds of M&E, such as M&E of community development or M&E of advocacy, where practical guidance is currently easier to find.

8. Conclusions

So is it really that difficult to carry out effective M&E of capacity building or capacity development? The answer is simple; both yes and no.

There are many examples of organisations that carry out effective M&E that enables them to build up a picture of individual or organisational change and learn in the process. There are also many examples of organisations that are able to illustrate wider changes resulting from improved capacity. In some circumstances this is easier than others. M&E is arguably easier in areas such as technical capacity building for clearly defined ends. In these cases the contribution of capacity building to end-results can be assessed by working forwards (to see the immediate results of capacity building efforts) or backwards (to assess the contribution of changed capacity to longer-term results).

For more general capacity building the challenges are greater. Here the effectiveness of M&E depends on a wide variety of factors. The evidence presented within this paper suggests that there are a number of key areas that need to be addressed in order to maximise the effectiveness of M&E.

- Be clear about the purpose of capacity building. Capacity building providers need to have a clear, stated rationale for carrying out capacity building, and a clear idea of what they want to achieve, both in the medium- and long-term. This might mean developing an appropriate theory of change. At the least it should involve developing clear, 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 the same as M&E designed to learn and improve; and there is little point in pretending otherwise. The purpose(s) for which M&E is carried out will have a large degree of influence over the types of approaches and methodologies used.

- Decide on the direction of M&E. Where it is important to highlight specific capacity building interventions then it may be more useful to attempt to evaluate the intervention itself, and work upwards (or outwards) to trace the results at different levels (or ripples). Where there are multiple interventions spread out over time then it may be more useful to start by trying to evaluate change at individual or organisational (or even societal) level, and work backwards to identify the contributions to those changes.

- Decide how far you intend to measure change. It is important to distinguish between changes that can be measured, and changes that can only be illustrated. Developing valid plausible links between measurable changes and wider goals may help enable M&E to be more realistic and less onerous in terms of time and resources.

- Use a sensible blend of tools, methodologies and approaches that will help provide a picture of what is changing (or not) and why. Where resources permit, findings should be triangulated by involving different stakeholders in M&E processes.

- Carry out M&E alongside capacity building support. Where possible, capacity building providers should make sure that any M&E processes are consistent with the capacity building process itself. This will help ensure that M&E supports the capacity development process rather than undermining it. It will also help to keep the costs of M&E down.

- If a donor is involved, agree key issues beforehand. This might include coming to an agreement about how far M&E should go in terms of measurement, and at what levels. It might also involve

agreeing the specific blend of qualitative and quantitative information required. Wherever possible, agreements should be recorded to reduce the risks of changing demands with changing personnel.

- Fight the battles that are worth fighting. In the current economic 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. At some stage there will be a need to produce some numbers that can show the scale or breadth of changes across different organisations. In most cases it will be 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 is often 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 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.

The evidence from this research is that the organisations that have been most successful in monitoring and evaluating capacity building work are those that have effectively addressed these key areas. But they are necessary, rather than sufficient, conditions. Indeed, they are not new, and similar conclusions have been drawn in many different papers and books over the past two decades. So the question still remains: what can be done to improve the overall quality and scope of M&E in this important area?

Impetus needs to come either from within or without. It has long been acknowledged that the effectiveness of any kind of M&E is heavily dependent on the interest, buy-in and commitment of senior management within an organisation. Again, one of the findings of this research is that many international organisations contain pockets of activity around the M&E of capacity building that result from the interests of staff at different levels such as regional or country field offices. However, the danger is that this will always be subject to change as personnel and interests change. This will leave us in the same position as before, with good work being carried out in isolated and fragmented cases; old initiatives dying out as new ones are developed.

Perhaps, then, the impetus needs to come from the outside. So what incentives could be introduced to achieve higher quality and more consistent M&E? Donors could respond in one of two ways. The most likely in the current climate (and indeed the evidence of history suggests that this is the most likely scenario in any case) is that further pressure could be applied to report on results. Bearing in mind the slow pace of organisational and societal change set against the short-term nature of much donor funding, and the distortions that can be introduced into M&E when it is linked to funding, this is likely to be a dangerous approach. At worst it will completely undermine learning approaches, and further the current tendency to reward capacity building providers that are most effective at using M&E for marketing or public relations purposes.

The second possible response would be for donors to provide incentives for capacity building providers that are willing to invest seriously in M&E for learning purposes in order to improve performance both within their own organisations and more generally across the capacity building community. The idea of making organisations accountable through learning is not new, and some might regard it as rather unrealistic. At the least it would involve some serious readjustment on the part of many donors. But if we acknowledge that little progress has been made in the M&E of capacity building over the past two decades, and that it is time for the whole area to receive an injection of added impetus, it is surely worth a try.

Annex 1: Bibliography

- AKDN (2009). *Aga Khan Development Network Civil Society Programme: Not a project but an approach to the whole civil society sector.* PowerPoint presentation developed by AKDN.
- AusAid (2006). *A Staged Approach to Assess, Plan and Monitor Capacity Building.*
- Bakewell, O., Adams, J and Pratt, B (2003). *Sharpening the Development Process: A practical guide to monitoring and evaluation.* INTRAC, UK.
- Barefoot Collective (2009). *The Barefoot Guide to Working with Organisations and Social Change.*
- CDRA (2001). *Measuring Development: Holding infinity.* CDRA, Cape Town.
- Chambers, R., Karlan, D., Ravallion, M. And Rogers, P. *(2009).* "So that the Poor Count More: Using participatory methods for impact evaluation" in *International Initiative for Impact Evaluation, working paper 4.*
- Christoplos, I., Mitchell, Y., Minear, L. And Wiles, P. (2005). *ALNAP Review of Humanitarian Action in 2004: Capacity building.* ODI, London.
- CIVICUS (2009). *Civil Society Index (CSI).* Downloaded from www.civicus.org/csi
- Cracknell, B. (2000). *Evaluating Development Aid: Issues, problems and solutions.* Sage, London.
- David, R. (1998). *Monitoring and Evaluating Advocacy Work.* ActionAid.
- Davies, R. And Dart, J. (2005). *The 'Most Significant Change' (MSC) Technique: A guide to its use.* April 2005.
- De Lange (undated). *Evaluating capacity development support: A gateway for capacity development.*
- Eade, D. (1997). *Capacity Building: An approach to people-centred development.* Oxfam, UK.
- 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.
- ECDPM (2008). *Capacity Change and Performance: Insights and implications for development cooperation.* Policy Management Brief 21.
- 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.
- Fowler, A. et al. (1995). *Participatory Self Assessment of NGO Capacity: INTRAC Occasional Paper Series no. 10,* by Alan Fowler, with Liz Goold and Rick James, INTRAC, UK.
- Garbutt, A. and Bakewell O. (2005). *The Use and Abuse of the Logical Framework Approach.* INTRAC, UK.
- Gosling, L. And Edwards, M. (1995). *Toolkits: A practical guide to assessment, monitoring, review and evaluation.* Save the Children, UK.
- James, R (undated). *Umoyo M&E of Capacity Building: Final report.* INTRAC, UK.
- James, R. and Hailey, J. (2007). *Capacity Building for NGOs: Making it work,* INTRAC, UK.
- 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.
- Mebrahtu, E., Pratt, B. and Lonnqvist, L. (2007). *Rethinking Monitoring and Evaluation.* INTRAC, UK.
- OECD (2006). *The Challenge of Capacity Development: Working towards good practice,* OECD DAC Network on Governance.
- Ortiz, A. and Taylor, P. (2008). *Learning Purposefully in Capacity Development: Why, what and when to measure?* An opinion paper prepared for IIEP, July 2008.
- PACT (undated). *Organizational Capacity Assessment Tool (OCAT) for NGOs.* PACT inc.
- Pearson, J. (2009). *Integrating Learning into Organisational Capacity Development of Cambodian NGOs.* Forthcoming article for Development in Practice.
- Phlix, G. And Kasumba, G. (2009). *Evaluation of Dutch Support to Capacity Development: Evidence based case studies.* ACE Europe.
- PSO (2004). *Monitoring and Evaluation of Capacity Building: Policy and instruments: a PSO manual.* PSO, The Hague.
- One World Trust (2008). *Global Accountability Report Indicators.* www.oneworldtrust.org.
- Reeler, D. (2007). *A Theory of Social Change and Implications for Practice, Planning, Monitoring and Evaluation.* Community Development Resource Association (CDRA). South Africa.
- Simister, Nigel (2000). *Laying the Foundations: The role of data collection in the monitoring systems of development NGOs.* Occasional Paper, Centre for Development Studies, University of Bath, UK.
- Stevens, B. (undated). *Measuring Capacity-development Results: In organisational development projects with groups of small to medium sized CSOs.* Songes Belgium, Brussels, Belgium.
- Uphoff, N (1991). "A field methodology for participatory self-evaluation" in *Community Development Journal 26(4),* pp271-285.
- VBNK (2009). *Annual Impact Assessment Report.* VBNK, September 2009
- Watson, D. (2006). *Monitoring and Evaluation of Capacity and Capacity Development: A theme paper prepared for the study 'capacity, change and performance'.* ECDPM, April 2006.
- 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.
- Venture Philanthropy Partners (undated). *Capacity Assessment Grid.* www.venturepp.org

Annex 2: Acknowledgements

Thanks to the following people who were contacted as part of the research, or supplied materials used during the research.

Laurie Adams (ActionAid International), Dr. Kaustuv Bandyopadhyay (PRIA), Brenda Bucheli (Independent Consultant), Nicola Chevis (VSO), Aroma Dutta (PRIP Trust), Jonathan Flower (World Vision International), Louisa Gosling (Independent Consultant), Volker Hauck (ECDPM), Simon Heap (Plan International), Richard Holloway (Aga Khan Development Network), Edo Huygens (Oxfam Solidarite Belgium), Daniel Jones (Christian Aid), Mark Keen (IOD), Robert Lloyd (One World Trust), Joyce Mataya (CABUNGO), Arjen Mulder (NOVIB), Christine Mylks (VSO), Steve Nally (DFID), Jonathan Patrick (DFID), Jenny Pearson (Independent Consultant), Jonathan Potter (People in Aid), Daniel Sershen (The Open Society Institute and Soros Foundations Network), Mary Sue Smiaroski (Oxfam International), Terry Smutylo (Independent Consultant), Sue Soal (CDRA), Dr Graeme Storer (VBNK), James Taylor (CDRA), Alix Tiernan (Christian Aid Ireland), Duncan Trotter (Save the Children UK), Yuko Yoneda (ActionAid)

Annex 3: Tools used for organisational assessment

This annex describes a sample of different tools that have been used for organisational assessment by different organisations in different circumstances. Not all of them would be described as an OA tool, but all would include common features that allow for the ongoing monitoring and evaluation of capacity by different groups. Most of these tools were submitted by different agencies contacted as part of the research.

Description	Breakdown of capacity	Rating/Ranking	System for ranking/rating	Analysis and action
The **McKinsey capacity assessment grid** is a tool designed to help non-profit organisations assess their organisational capacity. The framework and the descriptions in the grid were developed based on the input of many non-profit experts and practitioners (Venture Philanthropy Partners undated).	Capacity is divided into seven elements. These are aspirations, strategy, organisational skills, human resources, systems and infrastructure, organisational structure and culture. Each element is sub-divided into many constituent parts. The tool is largely based around internal aspects of capacity, rather than relational capacities or performance capacities.	For each constituent part four statements are designed to help score capacity on a scale of 1 to 4. Organisations select the text that best describes their current capacity in each area.	The grid is designed to be used as a survey. Organisations can sit together to decide on rankings, or can individually fill in the survey and then discuss.	The grid is designed to be adapted as required in order to identify those particular areas of capacity that are strongest and those that need improvement; measure changes in an organisation's capacity over time and draw out different views within an organisation.
The **Norman Uphoff Tool** was used in the People's Participation Programme (PPP) of the UN Food and Agriculture Organisation (FAO), which aimed to establish self-managed and reliant groups. The methodology described is designed to be a group's <u>own</u> method for strengthening <u>its</u> ability to meet its members needs through collective action (Uphoff 1991).	The methodology pre-defines 80 different areas (activities or modes of operation) which are included under six main headings. Groups are free to use whichever activities are relevant to them, or to add new ones. The headings are group operation and management, economic performance, technical operation and management, financial operation and management, group institutionalisation and self-reliance and other considerations.	Four alternatives are provided for every area. • the most satisfactory situation, with little or no room for improvement • satisfactory with some room for improvement • unsatisfactory with considerable room for improvement • unsatisfactory with great room for improvement	The methodology encourages long, full, frank and open discussion to reach a consensus in each area.	The tool is meant to stimulate discussion and argument. It is meant to be self-educative, self-improving. It is also designed to enable higher levels of the programme to monitor progress.
The **staged capacity building model** was developed by AusAID as a methodology for planning and monitoring capacity building. It is designed to be used by AusAID advisers and counterpart staff. It is not an OA tool as such, as it only concentrates on areas in which an organisation is assisted (AusAID 2006)	There is no breakdown of capacity. Specific areas of support are specified by advisers working together with counterparts.	In each area of support, the supported organisation can be assessed as: • dependent • guided • assisted • independent.	Advisers, counterparts, other members of the organisation and facilitators jointly decide where they are on the ranking scale.	Based on the analysis, an action plan is developed, and targets are set so that people can judge what progress has been made. This is reviewed at regular intervals.
The **SAFE system** was used by the Umoyo Network as an externally supported self-assessment tool. It was developed based	Capacity is broken down into a number of areas. These range from shared vision and mission to systems	Partners are asked to respond to each question in the following way.	Scoring is done in different level groups during a workshop.	The scorings allow a baseline to be established and perceived changes to be

on a survey of over twenty other OA tools (James undated) but also based on Umoyo network partners' criteria for a healthy partner.	for planning, m&e and reporting to relations with different bodies and programme performance. Each area contains a number of statements, such as "implementation is guided by plans" or "managers receive regular monitoring information that assists decision-making".	• Strongly disagree • Disagree • Neutral • Agree • Strongly agree	Scores are averaged to come up with a composite score.	measured at a later date. The scores are also quantified to supply information against the main donor indicators.
Spider diagram of institutional maturity described in Gosling and Edwards (1995). This is a simple diagram for plotting organisational capacity in different areas.	The key areas are technical operation and management, financial operation and management, linkages and negotiating levels, learning and evaluation mechanisms, accountability, degree of autonomy, funding/economic performance and organisational operation and management.	• 0 = undesirable level; drastic improvement required • 1 = poor situation: much room for improvement • 2 = good situation: some room for improvement • 3 = ideal situation: little room for improvement	The guide mentions that each aspect of organisational change should be vigorously discussed during participatory monitoring meetings.	Not specified.
The **RAISA Organisational Assessment Tool** was designed by VSO's Regional AIDS Initiative in South Africa. It was designed to monitor and evaluate evolving organisational capacity in organisations supported by VSO.	The six areas of capacity are strengthening service delivery, managerial development, operational development, relational development, strengthening national frameworks and HIV&AIDS workplace policy.	Progress is relative, and is rated on an annual basis against the following scheme. 1-No or very little progress 2-Limited progress 3-Good progress 4-Excellent progress	Objectives and indicators are set in up to three of the areas of capacity. These objectives and indicators are revisited every year and an assessment is made of how far the organisation has progressed.	The assessments are used to help develop the next annual plan. Some collation of progress is also carried out to report to external donors.
The **PSO M&E system** for capacity building was designed to gain insight into results that are being achieved in various dimensions of capacity building; and to learn which strategies/activities/methodologies, under which circumstances, were best suited to achieving the desired results. The system applied to everyone using PSO resources to support local partners in their capacity. However, it was accepted that partners would have their own M&E system as well (PSO 2004).	Capacity was divided into; 1. Human resource development (management skills, technical skills and attitude and motivation). 2. Organisational development (strategy and policy, learning capacity, structure, systems, staff, management style, networking, culture, financial management and technical skills). 3. Institutional development (strategic harmonisation, operational harmonisation, learning capacity, external influence).	In each sub-area there was a 'quality' score, an 'attribution' score and a 'sustainability' score'. Scoring was on a sliding scale of 1-4.	Not specified – it was assumed that under some circumstances one person might give the score.	In an annual report, people were asked to comments on scores for quality, attribution and sustainability. The OA tool was designed to be repeated to assess progress across time
Pact Inc. originally used an **Organisational Capacity Assessment Tool** developed in	In one version of the tool, capacity is broken down into governance,	The ranking scheme is as follows:	The NGO which is interested in the use	Participants reflect on which sections (or sub-sections)

Ethiopia as its starting point. It has further developed versions of this tool (generically called OCAT) in Botswana, Madagascar, Angola, and Zambia and is continually applying and modifying this tool in other countries of the world.	management practices, human resources, financial resources, mission competence, external relations and sustainability. All of these are broken down into sub-questions, and each sub-question contains a number of separate statements.	1 – this issue needs urgent attention and improvement 2 – this issue needs attention and could be improved 3 – this issue needs to be further examined 4 – this issue is basically well-handled 5 – on this issue there is no need for further improvement	of the tool arranges a meeting for this purpose, and engages a facilitator who has some experience in the process of organisational capacity assessment. Different stakeholders score each statement independently. The scores are then brought together and averaged	have the lowest scores (i.e. signifying that they are the issues in which the greatest amount of improvement is needed by the organisation). They also debate any differences in the results between the different groups of participants. The exercise often results in an action plan, and can be repeated after 2-3 years to establish progress.
The **One World Trust** has a framework designed to assess the accountability of different organisations. This tool is applied to large Northern organisations as well as those based in the South (One World Trust 2008).	The Framework identifies four dimensions of accountability that enable organisations to manage and balance the needs and interests of internal and external stakeholders: transparency, participation, evaluation and complaints and response. Each dimension includes a number of different indicators, which are presented as statements (such as 'the organisation has a specific policy that guides its disclosure of information').	Scores for each indicator are either '0' or '1'. Scores are weighted using a complex methodology to find a composite score for each area.	The basic methodology includes interviews with stakeholders, a literature review, discussions with external experts, the use of secondary data, and initial documentation followed by feedback. However, at the end of the day the tool is an external assessment, and the ultimate judgement lies with One World Trust.	The different scores are used to feed into a Global Accountability Report. In addition, each organisation receives recommendations and has a meeting to discuss the findings. Some organisations decide to take action based on those findings.
Transparency International's service delivery in Africa programme uses an organisational assessment to assess the capacity of national chapters. Because of the nature of the organisation the capacity areas are different to other, more generic, OA tools.	TISDA's system of assessing institutional capacity divides capacity into: • vision, focus and relevance • responsiveness, representativeness and non-discrimination • independence and professionalism	1 = needs radical improvement 2 = needs much improvement 3 = needs some improvement 4 = needs no improvement	National chapters were given freedom to decide how to carry out the assessment. Some filled it in after discussion and consensus. Other countries completed it individually and	The analysis was used to develop an action plan for capacity building over the period of the programme. The process will be repeated at a later date to gauge progress.

	• transparency and accountability • capability of carrying out effective research and advocacy campaigns in liaison with other actors These areas are then subdivided into further indicators of capacity.		showed the range of scores for each area.	
Oxfam in Belgium's **grid of criteria** is used to measure the progress made by partner organisations in order to achieve specific results. The Grid is based on several different organisational assessment tools Oxfam had at its disposal whilst designing the programme.	There are a number of different 'indicators', which are then broken down into specific criteria. The indicators are the quality of the decision-making process, the quality of the implementation process, the quality of the follow-up process, structuring organisation and functioning, the quality of the gender approach, learning organisation, the quality of mainstreaming risk reduction management, the quality of relations with other actors in risk reduction management and institutional capacities in risk reduction management. Each of these areas includes a number of different criteria.	Criteria are ranked as poor, good or high. There are pre-defined statements for each criteria that help with the assessment.	Oxfam Belgium suggests that it is important to do a reflection exercise with the partner before registering the quantified data. The tool is thus intended to be used as a participatory assessment.	The grid is being used in dialogue with partner organisations and its use will be subject to a learning-exercise in the course of the current programme (2008-2010).
Plan's **Child Centred Community Development tool** is designed to assess both knowledge and application of child centred community development (CCCD) in Plan's programmatic work.	The different areas are broken down into understanding the rights of the child, non-discrimination and inclusion, including gender equality, roles and responsibilities of right holders and duty-bearers, partnerships, multi-level approach, participation (especially of children and youth), social mobilisation, advocacy and accountability. Each area then contains 2-3 key process indicators identifying the necessary capacities.	For each area, ranking is based on both knowledge and application. The knowledge rankings are: • There is a good understanding of the CCCD element • There is some understanding of the CCCD element • There is limited understanding of the CCCD element For application, the ranking is: • There is evidence that CCCD element is fully operational and integrated into all	Ranking is expected to be based on focus group discussions with a range of different stakeholders, in addition to evidence from a range of different programme documentations.	Not specified

		programmatic work • There is some evidence that the CCCD element is operational and integrated into all programmatic work • There is little or no evidence that the CCCD element is operational and integrated into all programmatic work		
WaterAid has developed an **equity and inclusion tool** that is part of a mapping process to find out how different aspects of equity and inclusion are currently understood and implemented in WaterAid. Part of this tool is based on a scoring system to highlight the capacity of different parts of the organisation. The other part is a more general questionnaire asking wider questions about current practice and plans for the future.	Equity and inclusion is divided into four main areas: political will / commitment; capacity and resources; organisational accountability and organisational culture and values. Each area is then sub-divided into between 7-12 statements.	Respondents are asked to indicate how far they agree with each statement, and then come up with a composite score for each area. The ranking system for each statement is as follows: • Strongly agree • Agree • Disagree • Strongly disagree • Not sure	The tool appears in the form of a questionnaire that is send to members of a virtual team working on equity and inclusion. This team is encouraged to discuss the areas with their colleagues, and then provide rankings and comment based on the discussions.	It is hoped that the tool will help highlight examples of good practice that can be shared more widely in the organisation, and areas of weakness that can be strengthened. A strategy for mainstreaming equity and inclusion will then be developed to build on the strengths and address the weaknesses. In addition, it is hoped that the process of answering the questions should also stimulate thought and discussion about the ways in which equity and inclusion can affect different aspects of WaterAid and its work.