EMERGING PATTERNS IN THE CAPACITY DEVELOPMENT PUZZLE

Why, what and when to measure?

Alfredo Ortiz and Peter Taylor

25-Jul-2008

Abstract

Many capacity development (CD) programs and processes aim at long-term sustainable change, which depends on seeing many smaller changes in at times almost invisible fields (rules, incentives, behaviours, power, coordination etc.). Yet, most evaluation processes of CD tend to focus on short-term outputs focused on clearly visible changes. This paper will offer some ideas on how to deal with this paradox, by examining how monitoring and evaluation (M&E) does, or could, make a difference to CD. It explores whether there is something different and unique about M&E of CD that isn't addressed by predominant methods and ways of thinking about M&E.

Contents

1. Introduction—What should Monitoring &Evaluation (M&E) tell us about Capacity Development (CD)?

The debates around monitoring and evaluation (M&E) of capacity development (CD) are inextricably linked to the larger debates around CD and M&E individually. To monitor and/or evaluate CD presumes having a strong idea as to what its ends are in order to be able to tell the story of whether "it" is happening or not. And the relevance of different methods of M&E is dependent on whether the desired outcomes of interventions are process or product driven, easily or less definable, more or less time restricted, etc. But before delving into M&E of CD, it is important to define what is meant when we say "capacity".

Capacity has been defined from the very broad, e.g. "'[C]apacity' is understood as the ability of people, organisations and society as a whole to manage their affairs successfully (OECD, 2006: 12)."; to the more narrow, e.g. "Capacity is the ability to carry out stated objectives (LaFond and Brown, 2003: 7, quoting Goodman, 1998)." It has been qualified with value judgements about its proper use, e.g. "Capacity represents the potential for using resources effectively and maintaining gains in performance with gradually reduced levels of external support (LaFond and Brown, 2003: 7); and it has been defined with minimalist simplicity, e.g. "Capacity is a potential state [of performance] (Baser and Morgan, 2008: 87)." It has been divided into "hard" capacities such as "infrastructure, technology, finances, and staffing (Horton et al., 2003: 23) and "soft" capacities, such as the "…human and organizational capacities, or social capital of the organization, including such things as management knowledge and skills, and organizational systems and procedures,…[including] management information systems, and procedures for planning and evaluation (2003: 163)." The "soft" capacities have been divided even further, between tangibles—i.e. the systems and processes mentioned above—and intangibles, which highlight the importance of an organization having the "ability to function as a resilient, strategic and autonomous entity (Kaplan, 1999: 20)," as well as having the *capabilities* to commit and engage, adapt and self renew, relate and attract, and balance diversity and coherence (Baser and Morgan, 2008, Morgan, 2006)[1]. Merriam-Webster's online dictionary interestingly defines one aspect of capacity from the perspective of power: "the facility or power to produce, perform, or deploy."[2]

Notions of capacity relate also to broader issues of human development. Hotly debated over decades, and with origins in the field of biology, development has been equated by many with global economic growth which would result in all peoples of the world achieving economic parity with those living in the "developed" nations. Over time, "human development" has, however, acquired more complex meanings. The UNDP states:

> *Human development is about much more than the rise or fall of national incomes. It is about creating an environment in which people can develop their full potential and lead productive, creative lives in accord with their needs and interests. People are the real wealth of nations. Development is thus about expanding the choices people have to lead lives that they value. And it is thus about much more than economic growth, which is only a means - if a very important one - of enlarging people 's choices... .…Fundamental to enlarging these choices is building human capabilities —the range of things that people can do or be in life. The most basic capabilities for human development are to lead long and healthy lives, to be knowledgeable, to have access to the resources needed for a decent standard of living and to be able to participate in the life of the community. Without these, many choices are simply not available, and many opportunities in life remain inaccessible (UNDP website, accessed 2006).*

---

[1] These capabilities, in addition to the capability "to carry out technical, service delivery and logistical tasks," form the (ECDPM) capabilities model. Also see the insightful table on adaptive capacities in Horton 2003, p28-29.
[2] See http://www.merriam-webster.com/dictionary/capacity

The UNDP statement goes on to note that human development is also a right, and that "*the goal is human freedom. And in pursuing capabilities and realizing rights, this freedom is vital. People must be free to exercise their choices and to participate in decision-making that affects their lives. Human development and human rights are mutually reinforcing, helping to secure the well-being and dignity of all people, building self-respect and the respect of others*".

A number of well-known global frameworks and initiatives aim to support this goal, including the Millennium Development Goals, the Kyoto Protocol, Education for All (EFA), Food for All, the UNESCO Decade for Sustainable Development, and others. These frameworks do not provide a guarantee of positive change, however, as evidenced by the slow or absence of progress towards some of the targets enshrined within them. Progress is complicated by a wide range of variables that influence the process of human development, regardless of the goals and targets that are set. These include *economic, social, political and environmental* factors. *Demographic and climate changes*, the emphasis on the need for *sustainable development*, the desire for *peace-building and mutual understanding* between peoples to alleviate conflict and violence, *equality based on gender*, *ethnicity and religious belief*, and even the striving for *"happiness". A*ll play a part in determining global development pathways, yet all are influenced to a greater or lesser extent by *capacity* (Taylor, 2008).

The multiple angles of capacity demonstrate both the richness and the daunting nature of the subject. Monitoring and evaluating are, fundamentally, about measurement, which we look to, to help decipher this complex puzzle of CD. In order to improve CD design, implementation, learning, performance and impact, what, how and when to measure needs to be defined. What can or should M&E contribute to help CD be better understood and more effective? For M&E to be useful for CD, it has to tell us something about what works, what doesn't and why that matters. It should help tell the story over time about the role capacity has in development processes and how increased capacities help lead to increased development impact and sustainability. If M&E includes processes for observation, analysis, learning and improvement (change), then it should help substantiate the following assumptions[3] about CD in the broad field of education, and indeed in other sectors:

a) M&E of CD should help us understand development and development programming better, and inform our learning and adaptive management processes.
b) M&E of CD should show whether CD processes directly or indirectly strengthen other development processes and guide them towards impact in the most thoughtful way possible.
c) M&E of CD should aid in showing whether CD processes are developing or strengthening capacities that result in more system and organizational readiness and ability—"response-ability", including development of reserve and emergency capacity.

In this paper we do distinguish between *monitoring* and *evaluation*. Monitoring for us implies a more ongoing measurement process and evaluation a more periodic measurement process. Throughout the paper, in addition to M&E, we will use the term "measurement" as a broad reference to M&E. In our view, measurement should not always be made against predetermined criteria, standards or baselines, however, because those measures may bias us towards primarily looking for things that are in agreement with or in contrast to that which has already been identified. Measurement should also allow us to see unintended, yet important outcomes.

Considering each of the three key assumptions in turn:

---

[3] These assumptions are briefly answered in this section and in more detail throughout the paper. The remaining sections of the paper are not structured on the basis of these three assumptions.

a) *M&E of CD should help us understand development and development programming better, and inform our learning and adaptive management processes..*

Observing thoughtfully how capacity develops—increases, recedes, take leaps, evolves,—over time is an exercise in seeing whether some of the most important preconditions for development are present or not. Utilizing the broad view of capacities described earlier, we can see if resources, systems, autonomous motivation, skills, relationships and leverage, resourcefulness and other capacities are adequately present to effectively support social, economic, environmental and integrated development processes. Capacity strengthening processes should provide lessons on the correlations between capacity and development, and M&E should help document and process those lessons. It should tell us much on what works in development, what works in capacity strengthening, and how and where we need to change our CD interventions to maximize the usefulness of capacity in catalyzing other development ends. It should also tell us whether the environment/ecosystem is helping or hindering capacity development efforts. Key elements to be measured here are **learning for development and adaptive management.**

> *"Embedded in CD are questions about learning and change. In most CD work there is an implied logic that predicts how an intervention will affect CD, and how CD might affect other important development results (health, poverty reduction, equity and so forth). (Lusthaus et al., 1999: 15)"*

b) M&E of CD should show whether CD processes directly or indirectly strengthen other development processes and guide them towards impact in the most thoughtful way possible.

Whether capacity is a means to development, and ends of development, or both[4], over time its ultimate purpose is to enhance or strengthen the ability to achieve other development ends. It attempts this in some cases by trying to improve, replicate or scale up other primary development programmatic interventions (improving service delivery in health or education via training for example). In other cases it does not aim to strengthen specific primary development processes themselves, but aims to enhance *the ability* of different actors to strengthen these processes in a more macro sense (explored further in the next question).

Better developed capacity should eventually have a positive effect on other development processes (many of which are focused more closely on development ends), and M&E processes should show the extent to which these development processes are better off over time *as a result of* capacity development interventions. In this sense, M& E is uniquely placed to help make the case for CD.

Additionally, M&E processes should measure the quality of the capacity strengthening process itself, including the extent to which CD practitioners[5] design, facilitate and improve their processes with open minds in tune with ongoing endogenous development and its participants and history. M&E also explores the degree of transparency and cost effectiveness of what practitioners are doing. Key elements to be measured here are **relevance, effectiveness (impact) and efficiency of capacity strengthening processes.**

c) M&E of CD should aid in showing whether CD processes are developing or strengthening capacities that result in more system and organizational readiness and ability—"response-ability", including development of reserve and emergency capacity

Individuals, organizations and systems need capacity well beyond what they use on specific projects or activities each day. A predominant focus on directly linking capacity development to immediate

---

[4] This paper strongly supports the assertion that capacity is both a means and an ends. See the section on "standing capacity" for a more detailed discussion on capacity as an ends.
[5] This applies regardless of whether the practitioners are local or external.

technical performance will only lead to a lack of preparedness for atypical situations—which occur throughout everyday life. In addition to improving technical performance, strengthening processes should help develop broader "standing capacities[6]", including key intangible qualities such as relationship leverage, program design capabilities, innovative culture, autonomous self-motivation, and agile, adaptive management (Kaplan, 1999, Morgan, 2006). M&E processes should help note the existence and utilization of these capacities and examine their relationship to individual, organizational and system resilience, readiness, and ability in responding to development challenges over time. Key elements to be measured here are *resilience, readiness and ability in responding to development challenges over time.*

M&E of CD is more about understanding what is *worth* measuring than about what *can* be measured simply by applying comprehensive M&E tools and indicators. It starts with the assumption that the end purpose of M&E is continual learning and improvement—*change.* This implies that improvement is possible, desirable, and necessary for finding the best way to complement and strengthen development processes and individual, organizational and system response-ability over time. We assume also that capacity is for *performance,* but we recognise that changes in performance often take time (Kaplan, 1999: 10), and are not necessarily attributable to specific CD processes or approaches (James, 2001: 8). Standing capacity becomes critical and thus an end in its own right, since it is this which provides the readiness and reserves needed to respond to the unexpected, maintain basic functionality, and operate at high performance in the long term.

The remainder of this paper will explore in more detail the meaning of "capacity" and its links to performance and impact, in order to clarify what is useful to measure, when and why. The following section (2) will define capacity development ends in more detail in order to clarify what is being measured under the rubric "M&E of CD". Defining what we believe are the ends of CD will directly affect what should be measured and will allow us to determine if the unique nature of CD requires unique M&E processes. Section 3 continues with an overview of current M&E of CD dilemmas, focusing on what can be learned from key elements on the M&E of CD debate. Section 4 concludes the paper with some broad recommendations on moving forward.

We should note here that this paper does not set out to provide detailed or technical explanations on *how to measure* capacity development, although there is no doubt that this is a rich vein to explore, and which deserves further attention; not least by drawing together examples and experiences from the practice of different actors engaged in CD processes, and reflecting critically upon these. We do, however, provide some examples of how different methods may be used in relation to different types of evaluation for different purposes; these do have implications for approaches to measurement, and which we hope will generate further interest in continuing this field of discussion.

---

[6] "Standing capacities" are basic functionalities and unique organizational abilities—beyond those which are necessary for immediate performance—which are fundamental for long-term performance. This concept is developed in more detail in section 2.2.

2. Capacity development means and ends—"What are we measuring and when should we measure it?"

> "Over many decades billions of dollars have been invested in conducting projects, providing technical assistance to the countries where the project was developed, and educating and training high-level specialists in developing countries; yet lack of capacities to design, plan and implement policies and reforms remains today arguably one of the most serious obstacles to the implementation of EFA in developing countries. (Caillods and de Grauwe, 2006: 1)"

So begins "Capacity development in educational planning and management: IIEP's experience." The analysis in that document forms part of IIEP's current efforts to better define its CD strategies and processes. The capacity challenges noted, and the basic concepts that underpin thinking and practice of CD in response to these challenges, are applicable in many different fields, including education, which has historically been considered one of the principle ways of helping to build capacity. But as in other sectors, education policy, systems, processes and structures need to be examined to reveal the extent to which they have really grasped the essence of M&E for CD and whether they have been adapted and improved as a consequence. An examination of the role of performance and impact in guiding M&E of CD is our starting point.

> *"[C]apacity development is IIEP's core function...It is achieved through training of individuals, training of teams in the field, organizational development, and the promotion of appropriate policy environments. Research and networking also emphasize capacity development. Indeed, almost everything that IIEP does–be it in training, technical assistance, research, or dissemination–can be described as a form of capacity development. (IIEP Medium-Term Plan 2008-2013)"*

2.1. In search of performance and impact

The ECDPM capabilities model noted earlier balances out its broad mix of soft, hard and intangible capabilities (i.e. the capabilities to commit and engage, adapt and self renew, relate and attract, and balance diversity and coherence), with an additional core capability—"to carry out technical, service delivery and logistical tasks" (Baser and Morgan, 2008: 30). An important link is made to the technical or programmatic activities that organizations carry out to directly achieve their missions. This model, along with all of the literature reviewed for this paper supports the assertion that the "capacity for what" question has an answer, and the answer is *performance*, i.e. an organization doing what its mission or mandate says it does—and ultimately achieving impact. If this is true, for M&E of CD the most relevant questions regarding "capacity" then become:

a) Is capacity as a broad concept useful in determining what to measure (on the path to performance and impact)? Is it helpful to break down the broad concept of capacity in order to measure specific dimensions of capacity?; and,
b) How geared towards **_immediate_** performance should capacity and capacity interventions be?

a) Is capacity as a broad concept useful in determining what to measure? Is it helpful to break down the broad concept of capacity in order to measure specific dimensions of capacity?

We have seen that capacity, in its broadest sense, can potentially include everything an organization has at its disposal—monetary and physical resources, people, know-how, confidence, unique value proposition, positioning, relationships, etc.—for carrying out its work. It is capacity if it can be drawn upon, leveraged or put to good use—whether it originates from within or outside the organization. It has hard and soft, tangible, and intangible drivers that can be measured and can tell us something about performance and development over time. But does this broad understanding of capacity help us

to identify specifically how M&E does, or could, make a difference to CD at different levels? Or does, as Morgan points out, it leave us with an amorphous starting point?

> "At this point, capacity development becomes almost synonymous with development itself, making it difficult to assess results in a systematic way (1997: 2)."

Capacity as a "potential state of performance (Baser and Morgan, 2008: 87)" moves us one step further in that it describes capacity as a latent state whose potential energy is intended for use in performance. But what exactly should be measured? Should we be measuring positive changes in latent capacity, the application of that capacity, or the results (outcomes and impacts) that the application yields—or a combination of all three? If capacity is only a means to an end then can we not just focus on the ends and know that their accomplishment is a proxy for the pre-existence of capacity (LaFond and Brown, 2003: 8)? If capacity is an end in its own right then how should it be measured, and should measurement attempt to include precise references to longer-term development ends, or intended impacts? If those outcomes and impacts are impossible to attribute to a particular organization, funder or time-bound set of activities is it even useful to measure outputs, capacities and other predecessors to impact?

Baser and Morgan's more complete definition expands on capacity as simply "a potential state of performance" by identifying types and levels of capacity:

> "It is that emergent combination of individual competencies, collective capabilities, assets and relationships that enables a human system to create value (2008: 3)."

This definition includes *levels* of capacity (individual, collective, system), *types* of capacity (competencies, capabilities, assets, relationships) and *purpose* (enabling human systems to create value). Thinking about measuring the development of individual competencies, or the provisioning of tangible assets, or even the strengthening of collective organizational capabilities, seems less daunting than thinking about measuring "capacity". The more micro and specific we go, e.g. the capacity to implement procurement procedures, the more straightforward the task of M&E potentially becomes.

The following table (1) illustrates different types of breakdowns by capacity development purpose and shows how different types of capacity development interventions have different M&E implications[7]:

---

[7] This table is not intended to be comprehensive and aims only to draw connections between different areas of capacity and M&E.

Table 1—CD interventions, relations and measures

Purpose of CD Intervention	Intervention examples	Complexity factor[8] on achieving impact	Risks & relations to other capacities[9]	Measures
Broadly strengthen organizational capacities and increase organizational learning	• Processes that ask very broadly what the capacities and characteristics are of an ideal version of the organization, assess the organization against the ideal, and develop a strengthening plan to address the gap. Sometimes includes coaching. • One example is the DOSA[10] self assessment tool and process, which analyzes human resource mgmt, financial resource mgmt, service delivery, external relations, strategic mgmt & organizational learning (VanSant, 2000).	• High in general because of the tenuous links between broad organizational capacities and development ends. Extremely high if the organization doesn't have professional accompaniment for relevant weak capacity areas in which it lacks the internal ability to improve[11] (e.g. a specialized content driven area such as financial management, marketing, or a mission-related technical area).	• By design, takes a broad look at multiple capacities and multiple levels (although many overall tools are much narrower). The main risk is that so much area is covered that depth of execution will likely be superficial for some key capacities, or will fail to distinguish key capacities related to performance. Also, technical capacities are often given "equal" coverage with other capacities although they might be strategically more important. • How each capacity area relates to end, mission-level impact, is rarely tracked. • A key benefit is the learning orientation that helps personnel from different horizontal and vertical layers see themselves as part of a broad organization.	• Depends on the methodology. Usually a mix of organizational development (internal process changes), service delivery, and intangible (e.g. strategic management, or partnership leverage) changes /improvements. Outputs are tracked against the plan, and outcomes can become evident (via introspective discussion) if the organization periodically reassesses itself.
Strengthen organizational technical capacities	• Training of individuals in particular techniques. For example training a teacher in a new lesson planning process or a program technician in business planning for microenterprises	• High, and very difficult to find a direct correlation between the use of the new processes and higher level impacts (e.g. better educated students or viable microenterprises	• Organizations that over focus on technical capacities without strengthening the organizational systems and processes that support technical execution.[12] • On a practical level risks are high because of potential transfer problems between the individual and the organization (assuming	• Outcomes that show the competent application of the training, and perhaps inferences on the relation to impact (e.g. drawn from qualitative methods such as stories)

---

[8] This refers to the extent to which the desired impact potentially has multiple internal and external factors affecting it. The higher the factor the lower the amount of assurance that the CD interventions will take hold, or if they do, that they will be clearly attributable.

[9] By risks and relations we mean to identify the relationship between this particular area of capacity and others, and the risks that too narrow, or too broad of approaches might pose.

[10] DOSA stands for Discussion Oriented Self Assessment. It was created by Beryl Levinger of the Educational Development Center (EDC) and Evan Bloom of the Pact Capacity Building Services Group (CBSG).

[11] In theory the CD intervention would include processes to address the longer-term needs that the self-assessment identified. In practice, many self-assessments occur without adequate investment in the long-term implementation of the plan itself, often due to inadequate funding or short-term thinking.

[12] Our intention is not to suggest that all technical capacity development processes should be broadly presented with multiple support processes; rather we wish to highlight the relationship between capacities and how they affect each other over time.

Purpose of CD Intervention	Intervention examples	Complexity factor[8] on achieving impact	Risks & relations to other capacities[9]	Measures
		functioning)	the training was sound).	
Strengthen financial stability	• Training or coaching in fundraising, marketing, strategic communications, negotiation skills (e.g. with Ministries and other funders), etc.	• High. Multiple external factors—in addition to improved fundraising, marketing or communications techniques—affect whether an organization will be successful in improving financial stability	• If successful in increasing funding, organizations might confuse financial stability with organizational sustainability. Stability implies having enough money in the short- to mid-term to carry out an organization's planned activities, whereas sustainability implies having strong overall fundamentals in place that enable the organization to be a high-value asset over the long term. This touches on a wide range of tangible and intangible capacities. • If extremely successful, might lack the basic systems to manage it.	• Increased funding, diversification of funders, core (unrestricted) funding, alignment of funding with organizational core competencies and priorities (i.e. increased autonomy).
Strengthen a specific internal process (e.g. procurement)	• Process to map the purchasing cycle and document procedures for different types of purchases (e.g. products and services) and different threshold/risk levels.	• Medium to low—mostly internal organizational issues that could affect successful implementation.	• Most purchasing processes don't analyze cost-benefit in a macro sense (as long as it's in "the budget" the legitimacy of the purchase is accepted).	• Procurement policies and procedures exist and are consistently followed (applied). • Note: The value of trying to link to end impact is questionable.
Strengthen relational and adaptive management capabilities	• Participatory processes that link actors within a system and build or make visible areas of common ground. • Processes that help managers systemically analyze decisions, improve information for decision making and strengthen scenario planning	• High for both, multiple factors outside the control of the organizations	• Highly dependent on individual talents and motivations, many of which are difficult to cultivate if they don't already exist (Buckingham and Coffman, 1999).	• Very difficult to measure conventionally and outcome measures (e.g. participation in networks; better-informed decisions made) are potentially meaningless without an accompanying story or narrative.

Breakdowns of "capacity" are helpful because they are more understandable and relevant at an intervention level, where they logically lead us to the types of measures which will indicate their effectiveness. At this level they become actionable—the capacity for *what* becomes clearer and more intentional. Looking from too micro a standpoint we run the risk, however, of falling back into linear, direct cause and effect thinking or of not learning broadly about the links between capacity and performance, which should in turn inform an organization's learning and adaptive management processes.

For practical reasons we believe that a balance must be struck, that allows organizations to measure capacity in the most useful way possible without having to know in detail all that capacity is or is not. Understanding one's intentions with capacity development can be an important first step in analyzing capacity development complexities, but grounded in the reality of actual interventions within a particular context.

b) How geared towards __immediate__ performance should capacity and capacity interventions be?

Much of the literature on the subject shows a strong tendency to attempt to link the intentionality of CD activities to shorter term, immediate performance:

> "This debate about capacity as a means or an end of development generates little interest. Indeed, it attracts some disdain among many analysts and practitioners. It is usually viewed as somewhat irrelevant given the obvious need to deliver the 'results' in the short term, upon which many country participants and against which all development agencies are now judged (Baser and Morgan, 2008: 87)."

> "Capacity can only be defined in terms of a specific objective or goal. In the health sector, capacity does not exist for its own sake. Health planners and managers are concerned with capacity because it enables performance (LaFond and Brown, 2003: 8)."

> "…[USAID] mission programs usually view institutional capacity as a means to achieve higher level program results, rather than as an end in itself (Lessik and Michener, 2000: 3)."

> "Findings from studies on capacity interventions are generally based on qualitative studies and formulated in terms of performance changes. Attempts to evaluate CD also tend to take an instrumental view, looking at outcomes and impacts from a project/programme intervention view (Taylor and Clarke, 2008: 14)."

Yet "the connections between capacity and performance are not always clear...Patterns of both capacity development and performance are uneven, with progress going at different speeds and different times. Investments in capacity can take days or even years to yield significant results (Baser and Morgan, 2008: 87)." This is due to several factors, including the reality that an organization's performance "is influenced by its capacity, by its internal environment, and by the external environment in which it operates (Horton et al., 2003: 20)[13]."

Performance as a concept/definition may or may not encourage short term and/or limited technical output thinking of capacity. Baser and Morgan define performance as "...the ways in which organizations or systems apply their capabilities in daily use...It is capacity in motion.(2008: 87)" LaFond and Brown define performance as "…a result or set of results that represent productivity and competence related to an established objective, goal or standard. (2003: 7)." And Horton, Alexaki et al. express "an organization's performance…in terms of four key indicators: effectiveness, efficiency, relevance, and financial sustainability (2003: 20)."

---

[13] Based on Lusthaus, Anderson, and Murphy (1995) and Lusthaus et al. (2002).

The difference between these views of performance is important. In the Baser & Morgan definition, performance is simply the application of capacity, "capacity in motion", and is value neutral with respect to the types of outputs it yields or the timeframe in which it produces them. Adding qualifiers to this definition one might say that "good" performance yields outputs that show a clear trajectory towards ultimate goals and "bad" performance is tantamount to an organization exerting itself but not moving forward. For measurement purposes it would be important to know whether CD processes were leading to "good performance" over time. If performance is understood as capacity in motion, then it becomes neither desirable nor undesirable until we see where and how it is leading us.

LaFond and Brown's definition implies "good" performance, but then equates that performance with a result or set of results. This definition—leaping from capacity to results—seems to eliminate the time and distance between the application of capacity and the results that that may or may not yield. It potentially gets us back into the trap of expecting a short, direct path between capacity and results, and appears to focus on CD that produces harder, more tangible "productive" results.

Additionally, if performance is only seen in relation to an established objective, goal or standard, what does this say about organizational outputs that are more emergent and unexpected—and therefore unrelated to *established* objectives? Whereas pre-programmed performance might be a good indicator of capacity development in some cases, ***outputs and outcomes that are the result of emergent adaptive management/agile responses to complex environments are even more important proxies for capacity development.*** This is because they show evidence of an organization's ability to draw on learning-informed intuition and multiple intangible capacities—maintaining relevance in environments that are not perfectly predictable. They highlight organizations that are conscious of their place in their ecosystems and can thoughtfully represent constituencies and respond to the unpredictable nature of development. They may be harder to tie to immediate performance but they tell a fuller story of the correlations between capacity and development.

Performance is important, and we have argued that CD is ultimately for performance. But most capacities have only weak links with immediate performance (which isn't inherently a good or bad thing), and forcing these linkages in M&E plans only sets us up for failure when it comes to the time for measurement.

2.2.   Standing capacity

The idea of "standing capacity" mentioned earlier, challenges the notion that capacity and capacity interventions should be predominantly geared towards *immediate* performance. Individuals, organizations and systems need capacity well above that which they use on specific projects each day. This concept is easy to understand when it comes to readiness or lack thereof for natural disasters, where lack of standing capacities quickly reveals woefully inadequate responses that affect thousands of lives. But it also applies to an organization's readiness and ability to respond to routine and non-routine situations in general. In the context of education, routine situations might involve CD in the area of teacher training, for example, which is seen as a long-term investment in enhancing access, quality and institutional strengthening in schooling. If a region is devastated by an earthquake, however, and schools, and consequently the local education system, have collapsed, how well developed is the capacity of the system to help reintroduce schooling in order to avoid potential disruption of educational provision for a whole generation of school-age children?

We divide standing capacity into two areas: 1) basic functionalities; and 2) organizational talents[14], both of which are critical for good performance. Basic functionalities are the minimal systems,

---

[14] We use the term "organizational talents" as an adaptation of the definition of *individual* talents from the book "First Break all the Rules", which distinguishes individual knowledge, skills and talents as follows (our paraphrasing):
- Knowledge—what one knows
- Skills—What one know how to do

infrastructure, resources, collective ability and coherence needed for an organization to perform consistently well over time. They are about an organization and/or system's ability to be ready and able to respond to the range of *logical and probable* circumstances that normally present themselves. Organisations that are under resourced, overcommitted, poorly managed, donor driven or misplaced (e.g. an education institution that finds itself manufacturing local products which take it outside its core competencies, and negatively affecting the quality of its education) often lack this basic level of standing capacity. There is a level of standing capacity below which an organization simply cannot perform consistently well over time. "Basic functionalities" implicitly include a reasonable amount of "excess" capacity to be able to comfortably respond to every day needs[15].

The second level of standing capacity is what we refer to as "organizational talents". Organizational talents represent an organization's ability to summon, draw upon or leverage a unique combination of capabilities, resources, synergies, intuitions, relationships, etc., that allow it to be ready and able to modify plans, react, create, summon, innovate and be constantly relevant in the face of uncertain waters, *as well as* in routine situations[16]. Organizational talents are about high-performing, creative, innovative organizations whose synergies are well beyond the sum of their individual parts. Many of these talents come from a learned appreciation of being immersed over time in complex environments. Such capacities may be difficult to "build" through short-term interventions—organizations must often "live it and learn it"—but they are fundamental for organizational and system readiness and ability.

> "An understanding of capacity must also go beyond the instrumental, the technical and the functional and encompass the human, the emotional, the political, the cultural and the psychological. We can see these aspects of capacity at work in some of the cases. Some organizations lacked technical mastery in certain key areas such as financial management or project management. But they displayed enormous reserves of capacity in the form of collective resilience, social energy, courage, loyalty and ingenuity. These qualities enabled them to persevere and improve over time. (Morgan, 2006: 18)"

These intangibles are not just lofty principles or abstract concepts; having capacity, even in its latent stage, is being prepared to act. If organizations only prepare for their limited results and immediate programming, then they are not preparing systemically. M&E of CD needs to measure standing capacity in order to better understand the links between capacity and performance[17].

Standing capacity is important at a systems level as well. On a macro level if development is expected to result in a broadening of freedoms (Sen, 1999), then the capacity of a system, with its

---

- Talent—What one does, thinks or feels consistently (well) and can put to productive use

"A recurrent pattern of thought, feeling or behaviour that can be productively applied. BUCKINGHAM, M. & COFFMAN, C. (1999) *First, Break all the Rules: What the World's Greatest Managers do Differently*, Simon and Schuster."

[15] Some sports provide a good metaphor for reasonable excess capacity. In professional basketball for example, although only 5 players play at a time, and very few teams rotate in more than 10 players during a game, professional teams usually roster 12 players to provide adequate backup for injuries, practices and other unforeseen circumstances. Public sector work, including education, is often perceived as suffering from inertia, and being unresponsive and inflexible. This is not necessarily caused by a lack of capacity of individuals to undertake their work, however, but may result from an absence of systems that support greater responsiveness of employees. Excess capacity may therefore be supported by decentralising policies, which encourage greater autonomy and more flexible responses – such as hiring extra staff when needed or by providing incentives to those already there to undertake additional tasks and responsibilities. .

[16] The ECDPM capabilities are applicable at different levels, both for basic functionalities and organizational talents. The more intangible elements of that model are particularly relevant to what we are describing as organizational talents.

[17] It should be noted that many of these intangibles are addressed in different capacity assessment processes and tools (see VanSant, 2000), but their centrality as enablers of dealing with complex systems might be diluted when mixed in with discussions on procurement systems and other harder capacities.

actors, incentives, norms, processes etc. to be able to broaden these freedoms is paramount. Investing in the capacity of the health sector (public and private actors) *to be ready and able* to respond to big and small, recurrent and emergent health challenges, and to guarantee people the freedom of good health is an important investment. That capacity must be developed so that the system is ready and able, both for running individual clinics more efficiently today, as well as to be ready and able for a public health emergency that might never happen; or for a global pandemic that most likely will happen, but perhaps tomorrow or perhaps in 10 years (e.g. a public works project that builds levees prepared to handle a 50 year flood). Countries and local governments that invest in education and health capacity, without having precise control over how the capacity will be applied in the future, are strengthening system-level standing capacity.

Readiness and ability are preconditions to guaranteeing freedoms, even when those freedoms aren't immediately threatened, and even when not tied to an immediately identifiable sectoral outcome. There is an assumption that capacities will eventually be exercised, put to use, applied to reduce "unfreedoms". The ends of CD are for eventual performance, albeit often far into the future, but they have a present value that makes them worthwhile of investment today.

An organization, sector, region, network or other type of formal or informal system that is prepared (ready and able) to act is an end in itself and should be invested in and measured.[18] This belief is a core assumption behind many sector wide approaches, regional projects, governance and network strengthening processes. In the broadest sense "good governance" is in effect a form of system "standing capacity".

As a final note regarding system-level standing capacities, the concept used here includes the capacities of system level actors and formal governance processes, as well as the capacities of individual actors to both develop and apply their organizational talents individually, *and* become an active part of their ecosystem. Much of the literature highlights the importance of relationship capacities (Crutchfield and Grant, 2008, Morgan, 2006), and many other organizational "talents" implicitly or explicitly reference the connection between an organization—which is an open system— and its relation to its larger ecosystem with its complexity of processes, opportunities and challenges.

> "…High impact non-profits work with and through other organizations—and they have much more impact than if they acted alone (Crutchfield and Grant, 2008: 107) ."

> "Capacity has to do with collective ability, i.e. that combination of attributes that enables a system to perform, deliver value, establish relationships and to renew itself. Or put another way, the abilities that allow systems - individuals, groups, organizations, groups of organizations - to be able to do something with some sort of intention and with some sort of effectiveness and at some sort of scale over time (Morgan, 2006: 7)."

> "Capacity as a state or condition is inherently a systems phenomenon (ibid)."

As discussed earlier, M & E of CD is useful to the extent it helps change related behaviour—i.e. improves CD design, implementation and learning—and/or improves accountability to relevant stakeholders. By first understanding the intended purpose of capacity development interventions— and the underlying assumptions, relationships and complexities—we can get a better idea of what specifically we want M&E to help accomplish, and subsequently choose the best methodologies and tools for carrying that out.

---

[18] As referenced earlier, systems performing poorly in natural disasters. Additional examples include low capacity networks such as many protected area systems, low capacity governance systems such as a regional government and the sub actors and systems within, and multiple other complex systems that are expected to function at a system level.

"[N]o simple recipes or blueprints are suitable for evaluating the broad range of organizational capacity development efforts that take place in different organizations. In one case, the organization may wish to evaluate a capacity development initiative that is just getting under way, to sharpen its goals, and consolidate its approaches. In another case, it may want to evaluate the results of a 'mature' or completed capacity development initiative, to report on impacts and benefits to its key stakeholders (Horton et al., 2003: 84)."

3. What can we learn from M & E of CD dilemmas?

Many donors and sub-donors clearly want to be able to show how "X" amount of resources invested resulted in "Y" amount of improvement, change or development. The broad field of education is no exception. This desire ranges from simply wanting to know if the investment strategy is working, to being able to report to back donors (i.e. legislatures, taxpayers in some cases) as to the efficacy of the use of funds. Whether it is capacity development of teachers to educate students effectively, or something as technical as the construction of school buildings or water wells, there is a desire to measure whether an investment is working and the extent to which a particular donor's investment is contributing to positive change *as they intended*.

The problem lies in the "as they intended". Donors and sub donors tend to create a tremendous pressure for those whom they support to plan and measure much more than they could possibly, intelligently predict. The consequences of this pressure are manifested through huge efforts by implementing actors to show results in order to comply with sanctioned programming mandates (e.g. staying within what a particular donor or development agency permits and supports with its funding), and to demonstrate a direct relationship between cause and effect. This situation lays the foundation for much of the linear thinking that is endemic in development programming (Kaplan, 1999: 12, Pasteur, 2006: 22). A matrix culture is created that attempts to show how resources, a plan and efficient project execution will result in delimited, predetermined change. "Despite the high levels of uncertainty within the development environment, there is a pressure to be able to predict and to appear infallible (Pasteur, 2006: 35)."

Examining some of these dilemmas further shows how they affect the way we understand, plan, execute and measure CD[19]. Four dilemmas are particularly relevant for M&E of CD:

a) Static development versus development in motion
b) Linear versus complex adaptive systems (CAS) thinking, programming and measurement
c) Attribution
d) Donor accounting focus versus open learning approaches

3.1.   Development being a process already in motion

Allan Kaplan in "The Developing of Capacity" starts off early on with the story of the eager Nikos in "Zorba the Greek" coaxing a butterfly into a premature birth, and ultimately lamenting its inevitable death (1999: 2). Much of the capacity development literature stresses the fact that development is already happening before any project, donor, program, initiative, etc., ever arrives, and to not recognize this is an irresponsible error and ultimately a precursor to an ineffective use of resources. Too many donors and executing agencies are determined that their projects be

> *"[D]evelopment does have a pace of its own. There is an absolute limit to the extent to which it can be speeded up through the application of increased resources and developmental interventions (Kaplan, 1999: 10)."*

---

[19] We are aware that even the way we speak (e.g. "…affect the way we understand, plan, execute and measure CD") is indicative of how prevalent externally driven CD processes are in many contexts.

executed in any event, and when those projects are severely out of tune with the development processes already in motion, they are likely to fail. They fail because:

- Capacity development programming that does not recognize development in motion is quite literally a foreign object; i.e. it pushes ideas that aren't likely to take hold because they are out of step with local realities;
- They do not build on momentum; i.e. positive development initiatives and processes already in motion; and,
- The motivation needed to take forward a strategy that does not fit will in turn require a *push strategy* to convince people to carry it out (Kaplan, 1999: 5). Even when the appropriate incentives are put into place, true motivation will be dubious because participation will likely be led by the possibility of short-term gain. The fundamentals required for sustainability will be lacking and therefore the project activities and desired behaviour changes are unlikely to develop deep roots.

M&E of CD should periodically test whether the selected interventions are likely to take hold based on their fit/relevance to ongoing development processes. They need to discover the extent to which CD processes are making an honest effort at thoughtfully "reading" [20] the situation before acting[21]. This requires serious time and resources spent studying a context before intervening, as well as in continually monitoring the relevance of selected approaches and interventions (Kaplan, 1999: 8). Participatory, asset-based approaches can give a basic reading of the moment in development that CD interventions hope to build on, but they will always be limited by partial understanding. A profound understanding of development is a daunting task even for those who are living it. Hence, the importance of ongoing monitoring and adjustment.

3.2. Linear versus complex adaptive systems (CAS) thinking, programming and measurement

Besides being unable to predetermine precise long-term outcomes without knowing where one is starting from, it is inadvisable to attempt to *precisely* predict these outcomes in any case, because multiple, unpredictable factors are often at play that influence the direction that development takes in practice. *Complexity theory* sheds light on the futility of setting up planning and measurement schemes that assume CD interventions have more control over their desired ends than they actually do[22]. One of the basic premises of complexity theory for CD is that *the directions in which "development" is going have little to do with where well-planned CD interventions intend for it to go.*

> "Complexity theory posits that it is not possible to predict with any confidence the relation between cause and effect. Change is emergent. History is largely unpredictable. Organised efforts to direct change confront the impossibility of our ever having a total understanding of all the sets of societal relationships that generate change and are in constant flux. New inter-relational processes are constantly being generated, which in turn may affect and change those already existing. Small 'butterfly' actions may have a major impact, and big ones may have very little impact. (Eyben et al., 2008: 203-4)"

---

[20] Kaplan, Morgan and multiple authors cite the importance of "reading" the context before intervening.
[21] This has very practical implications even at the intervention scheduling level. In 2007, while negotiating with a local organization the details of an organizational sustainability process that was being funded by another international NGO, the organization's director was adamant that we read the reports of previous CD processes (strengthening processes, strategic planning, etc.), before talking further about design. Furthermore, she insisted that we structure the workshop component of the processes in such as way as to minimize interrupting the work cycles. She was happy about the process, but felt past external processes had been *applied on* her organization in a pre-determined manner, not building on existing processes and momentum and not respecting the way the organization worked and learned best (intermittent workshops versus intense "lock-downs").
[22] See the column "Complexity factor on achieving impact" from Table 1.

In practice (and if looked at retrospectively) capacity development moves in directions where its trajectory and momentum are taking it (which will result in poorly placed CD interventions if not understood beforehand); and where other actors and influences—including policies, internal and external power structures, culture, weather—and other visible and invisible factors push it.

"Capacity is an emergent property that evolves partly through the pushes and pulls of contextual factors including global economic trends, national governance, the legacy of regional history and many others. The capacity of an organization derives much of its character from its interaction with the bigger systems within which it is embedded (Morgan, 2006: 18)."

"We may start with a set of expected activities and results, but as events unfold, the project or programme needs to respond, change its theory and revise its expectations. The emphasis should be on setting up systems for monitoring the impacts of our work, changing directions accordingly and monitoring those changes in direction. We need to constantly think how what we do relates to the goal, but we also have to hold that theory loosely and be prepared to learn and change.(Bakewell and Garbutt, 2005: 21)"

Capacity strengthening interventions only have a chance of affecting development to the extent that they're adequately in concert with ongoing development processes. Their effectiveness depends also on whether other actors and influences are cumulatively overwhelming the ability for specific interventions to promote change.

For M&E, especially of broad capacity development initiatives, this implies *using iterative planning and measurement approaches* that assume change is emergent, unpredictable and based on multiple factors. Bakewell and Garbutt offer some practical advice on how to operationalize this:

"Rather than tying ourselves into one overall model of how the project will work, which relies on impossible predictions, we could start by focusing on achieving consensus around the initial activities. Thus we agree on the first step on the way. As action is carried out it must be reviewed and the next steps are determined, in discussion with primary stakeholders. Under such an approach, implementing a project is not concerned with following the predetermined path. Instead, it is like feeling your way through the marshes, stopping regularly to reassess your situation, changing direction, possibly even retreating as required, in order to work through towards the goal without stepping off a solid path and becoming impossibly lost in the mud (see Figure 1) (2005: 20)."

Figure 1[23]: Feeling the way to the goal (ibid)

---

[23] In "A Theory of Social Change" (2007), Doug Reeler uses multiple diagrams that complement the Bakewell & Garbutt concept shown here (and develops it further by types of change). But worth noting is that in addition to the emergent pathways, the "desired future situation" is also dynamic/changing in Reeler's models—which we believe more accurately reflects emergent development—whereas it appears fixed in this diagram. An attractive feature of the model above, however, is the concept of developing initial consensus on an intentional short time frame and then continually re-planning as information and smart directions emerge.
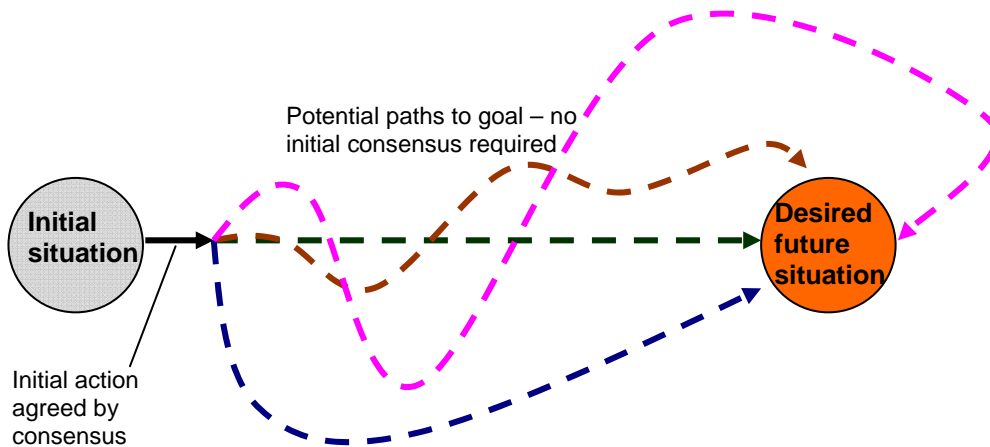
Currently, the project manager is seen as the lynchpin in many development processes, because he or she is given the task to keep different actors on track, ensure the planned activities take place, measure outputs and use resources efficiently and on time. We believe that a new type of guidance is needed for CD interventions, however, which means new roles, tasks and responsibilities, and different ways of understanding how change happens and how to measure change. This is true especially for development practitioners who are determined to "feel their way through the marshes (Bakewell and Garbutt, 2005: 20)", ask questions (Kaplan, 1999), and then emergently and creatively test capacity development interventions for relevance and potential for leading to development impact. If CD is a long-term process through uncharted waters, then more navigators and navigating processes (as opposed to simply training up more expert project managers) are needed to adjust course and keep moving forward over time. Although we are keen to find good news from M&E, and certainly donors and funders are hoping for this, we should be wary of the claims of any complex capacity development project that shows it has come impressively close to sticking to its original project plan and indicators. This may well be an indicator of project managers wearing blinders, and missing opportunities for more thoughtful programming over time.

Iterative measurement, which implies "stopping regularly to reassess your situation, changing direction, possibly even retreating as required," can include formal periodic evaluation, but is best served by ongoing monitoring processes. In these often-informal spaces, practitioner experience and intuition can be combined with program evidence (e.g. surveys, reports, etc.) in a decision-making setting where plans are debated and improved. The starting assumption should be that the goal of the process is not to close the gap between planned and executed activities (i.e. measurement against a static plan); rather, to reduce the gap between the plan and reality on the ground, including emerging opportunities and challenges. If the plan is sacrosanct, then the project participants become the tools, and opportunities for progress within the complexity of it all are forgone.

3.3.   Attribution

This dilemma is closely related to the previous one in that it also assumes that multiple factors affect eventual impacts, thus making it impossible to attribute most impacts to particular CD interventions:

> "Just as a ripple becomes smaller and more difficult to see the further out it goes, so it becomes more and more difficult to attribute any changes at beneficiary level to the original CB intervention. As Oakley asserts, 'as a project moves from inputs to effect, to impact, the influence of non- project factors becomes increasingly felt, thus making it more difficult for the indicators to measure change brought about by the project'(Oakley and James 1999:23). As one moves outwards the less control the original CB provider has on what happens (James, 2001: 8)."

"As a result it must be stressed that we are looking for plausible association, not direct attribution (ibid: 12)."

Attribution in education programmes has long been a particularly challenging issue. So many variables influence behaviour of individuals in society that attributing behavioural change to specific educational interventions (whether schooling, non-formal education or training) is known to be fraught with danger, and a major methodological challenge. Attribution in CD therefore has another interesting challenge—the broad debate discussed earlier on difficulty of establishing links between capacity, performance and end impact, whether at the individual or the organisational level:

"Interventions can happen at a variety of different levels, for example providing training courses for individual staff members, team building, mentoring for senior managers and visioning and strategic planning at an organisational level. These investments in organisational functions and processes would aim to result in an actual change in programme performance and, ultimately, in the lives of the poorest and most marginalised groups. This presents particular challenges for assessing impact, not least whether it is possible to demonstrate a causal link between a particular organisational intervention and a wider process of change. For example, can a link be found between establishing staff performance appraisal procedures and the resulting improvements in the lives of the most vulnerable (Hailey et al., 2005: 5)?"

Perhaps more important than allocating credit for different levels of impact is assembling evidence to show that one is learning diligently, adapting and taking well-informed paths.

"It doesn't really matter whether you can quantify [or attribute] your results. What matters is that you rigorously assemble *evidence*—quantitative or qualitative—to track your progress. If the evidence is primarily qualitative, think like a trial lawyer assembling the combined body of evidence. If the evidence is primarily quantitative, then think of yourself as a laboratory scientist assembling and assessing the data (Collins, 2005: 7)."

We believe that to some extent the attribution "dilemma" isn't really a dilemma at all. The fact that direct attribution has been so difficult to assign to most CD interventions and programs, is an indicator that we're living in non-linear, complex adaptive systems, which demand open learning approaches in order to navigate towards impact. The attribution dilemma is actually a useful clue to how capacity development works.

3.4. Donor accounting focus versus open learning approaches

"In practice, capacity-building M & E is often encouraged (or required) by external stakeholders to be used primarily for accountability (LaFond and Brown, 2003: 5)." This is a debate about M&E and its use in development in general. The argument, outlined already in this paper, is essentially that externally driven M & E primarily based on accounting[24] to donors, does not seem to yield learning, and actually may produce counterproductive behaviour. There is often a focus on predictable, linear causal chains in logical frameworks that cause organizations to overlook the complexity of the system within which they operate, and not be open to reading the situation and using adaptive management to find the best path forward:

---

[24] Most authors refer to this as "accountability" requirements, which we believe does a disservice to the broad concept of accountability which includes heavy responsibility to stakeholders as well. We suggest using the term "donor accounting" or "upward accountability" (GUIJT, I. (2007) Assessing and Learning for Social Change. IN IDS (Ed.) *IDS Discussion Papers.* IDS.) which more accurately describes the narrow concept which often is referred to as "accountability".

"Moreover, many approaches to impact assessment have tended to be mechanistic, linear, standardised and symbolic in nature. Furthermore they are often disconnected from real decision- making. Where sufficient data is generated it is often not analysed appropriately nor disseminated in a timely or user friendly way. This has provided little opportunity for meaningful reflection or for putting learning into practice to improve performance (Hailey et al., 2005: 5)."

M&E, within a narrow understanding of "donor accounting" is likely to be instrumental and focuses often on finding novel ways to satisfy the requirements laid down by funders, whilst often hiding what has actually taken place, if it has not fitted closely with pre-determined expectations. Time is then wasted on tracking outputs that say little about impact. Watson (2006) breaks down the themes into "two broad 'schools': those who have pursued results-based management approaches (embodied amongst others in the project framework), and those who advocate systems thinking- based approaches." He continues:

"the ECDPM case studies illustrate that sustainable development and change take time. However, results- based management approaches tend to stress short-term 'products' or delivery and tend to discourage the emergence of long-term processes of change unless they are carefully tailored to the context (Watson, 2006: vii)."

The implication of these and similar statements in the CD literature is that results based management (RBM) approaches are incompatible with systems thinking and learning, only marginally focused on impacts, and more focused on narrow donor accounting of resources and activities. Measuring impact "requires the use of a range of client- focused M&E approaches that go well beyond assessing of the attainment of delivery targets (ibid)."

We are demonstrating here our belief that rigid, inflexible and pre-determined project frameworks lead to a similar mode of M&E, which in turn limits learning. We could compare this situation to that of school students rote-learning in order to satisfy the questions of their examiners; they may "pass", but their learning is unlikely to contribute much to their capabilities to function well in society beyond utilising further their ability to satisfy the demands of others. We have not shown however, whether learning does emerge differently in M&E systems that are less rigid, more open, and not determined primarily by the need to account to donors. This is a challenging assumption. Much of the literature does not explore this theme and seems to assume that if organizations could only rid themselves of these questionably useful systems, then learning and emergent, adaptive management would occur. Although some authors note the substantial cost and effort needed to implement effective learning systems, e.g. (Hailey et al., 2005), very little is discussed of the cultural and practical shifts that are required to implement such a system.

Some argue, that what is needed is a mix of the two—a system or processes that effectively incorporate both donor accounting and internal learning needs , accountability and learning, although others insinuate that the two are inherently incompatible (Bakewell and Garbutt, 2005: 11, Earle, 2002: 7), and what perhaps should be considered is an "amicable divorce" (Mebrahtu et al., 2007: 139). One response would be for learning systems to be set up separately and complex M & E systems stripped down to the bare minimum of donor accounting (Earle, 2002). But whether separate or combined, we need to be clearer what we expect M&E learning approaches to CD to deliver.

What would we want from learning approaches to M&E of CD that donor accounting approaches cannot deliver?

An emerging trend in the literature of M & E of CD suggests that a broader "learning approach" to M&E of CD is better for learning and adaptive management and ultimately is better also for measuring impact, which is best "told" through a story than measured using predetermined indicators.

"How can development professionals gain a more nuanced understanding of the highly contextual and often ambiguous environments and relationships in which they are involved to make appropriate choices and decisions? More consistent and collaborative processes for holistic and profound reflection and learning are seen as fundamental to improving practice in this respect." (Pasteur, 2006: 22)

"Where measurement seeks to apprehend meaning, it has to be seen as a longer-term, endogenous and creative process. People's stories provide a critical window and reveal a great deal about the invisible forces that shape tangible phenomena." (Dlamini, 2007: 4)

We believe that there is real value in piecing together different kinds of evidence and telling stories about how development actors think change and development happens. Showing how these actors, whether individuals or organisations, have adapted continually along the way to reach a positive point seems more helpful than assembling a set of disjointed outputs that tell no story, and ultimately add up to very little.

This is probably shown best using an example. Let us imagine an organization (such as a teacher training college) going through a self-assessment process which aims to highlight areas for improvement through CD. Such a process might well be facilitated by an NGO (international or local) focused on CD service provision. Following normal practice, a CD action plan would be developed, as an output to the service provider (SP). In the following table we show in the left hand column some outcomes (again these are outcomes to the SP) that might have resulted from the execution of the action plan. Although a fictional example, we have seen many similar lists or outcomes. As a counterpoint, we then show in the right hand column how we might imagine using a story or narrative, to describe the same outcomes:

Outcomes	Story
Performance area: *Leverage from strategic alliances* • Monthly committee meetings are being held to further develop our alliances strategy • Increased participation in key conferences and networks	• *"Last week at a conference I met with some of the Principals of other Teacher Training Colleges in our State, and was able to fluently explain our unique approach to developing curriculum to address the needs of a diverse range of learners—whereas before I knew we were unique but couldn't explain it. And when I came back to the office I felt like I could better explain to my staff where to look for complementarities and opportunities with our peer organizations."*
Performance area: *Financial sustainability* • Developed criteria for selecting opportunities • Increased # of responses to requests for proposals (RFPs) (# of proposals submitted) • Carried out a "get out the word" workshop attended by 5 major funders in the education sector • Overall funding increased	• *"I can truly say that over the last year we have become more thoughtful and selective in the types of opportunities we pursue. In the application of the criteria for selecting funding opportunities we became painfully aware of how much we chase opportunities for survival—even in areas that are only marginally related to our core competencies. But about half way through the year we realized in a meeting that looking back, we had actually exercised some discipline and not pursued about 5 major opportunities that really didn't make sense for us. And the collaborative process we've put into place has helped our faculty think more strategically, and to be more selective in pursuing opportunities."*

Looking at these examples, where do we find the energy and meaning? Both examples show successful execution of strategy in the outcome column, but in the story column there seems to be greater recognition of the importance of allowing new things to emerge, and making good choices about which directions to take. There seems to be greater application of ***strategic thinking*** in the

story column—an intangible that is difficult to capture with indicators, but which is clearly important for long-term performance. As participants in development interpret and express how different capacities became meaningful (i.e. useful, relevant) and applied it becomes easier to see what matters and what leads to impact. And in the process their capacities for "evaluative thinking" are developed, which is fundamental for M&E of CD (Horton et al., 2003: vi., Watson, 2006: 26).[25]

The *Most Significant Change* (MSC) Methodology is particularly relevant in this aspect in that it collects significant change stories which infer some relationship between what an organization and relevant stakeholders feel signifies important change, and the multiple factors that contributed to that change. The methodology then involves rigorous vetting and validation processes to make sure that the evidence of significant changes is leading to understanding of impact and its drivers, and improvement of performance (Davies and Dart, 2005, Wrigley, 2006).

Table 2 shows four evaluation types with different focuses and uses for capacity development evaluation, including learning-focused approaches[26]. The question of rigour often arises, but we believe that the perceived rigour of these approaches is less important than their ability to create a dialogue on linkages between what CD interventions do and what they lead to. This approach allows for meaningful interpretation to guide adaptive management and improvement. In dynamic, complex systems, we sense that observation and study, learning, abstract framing, adaptive management, and agility in changing plans and putting learning to practice are more important than rigorous tracking of outputs that ultimately do not reflect at all the reality of the situation they are describing. The focus instead is on seeing and understanding the patterns and factors that came into play, thereby creating the conditions for change. We do not suggest that such approaches can convey "realities" fully, since these are constructed by those directly involved. They do have the potential, however, to support much deeper learning, in ways that are more collective and inclusive, and are more likely to lead to profound understanding and action, and therefore useful change.

In the table we also identify some possible methods that could be used to support learning within different evaluation types. Many of these methods are sophisticated and have been developed and tested in a range of different contexts; by combining different methods, it is possible to generate both quantitative and qualitative data, which helps to enrich the findings, and hence the learning process.

---

[25] Watson was referring to "Outcome Mapping" EARL, S., CARDEN, F. & SMUTLYLO, T. (2001) *Outcome mapping: building learning and reflection into development programs,* Ottawa, IDRC..

[26] It is worth noting that many learning based evaluation techniques are time and cost intensive. That said, "Practitioners are increasingly recognising that the benefits of using a variety of different approaches outweigh the costs. The value and credibility of these assessment systems however, depends on how well they can be adapted to local circumstances while still being internationally accepted and comparable (Hailey, et al, 2005)."

Table 2—Selected evaluation types and different focuses and uses for capacity development

Evaluation type	Purpose and assumptions	Focus	On learning	Possible methods[27]
General impact evaluation	• Tells us whether an organization is being effective—achieving its mission. • Impact is the ultimate test for capacity, i.e. a high impact organization is a high capacity organization. As such, this is the most direct measure of capacity. • A well run organization that isn't delivering on its mission is a low capacity organization.	Impact	• Although learning for CD is secondary, much can still discovered by taking impacts and working backwards to capacities to identify strengths and weaknesses and their connection to performance and impact (to see the extent to which capacities are being applied effectively).	• Impact evaluations • Outcome mapping
Stakeholder-centred impact evaluation	• Determine the relevance of organizational capacities from the perspective of key stakeholders and users of organizational services. • Parameters for usefulness and relevance of organizational capacities and their application are defined by the stakeholders themselves, and evaluations reveal the meaning that participants derived from the processes. • Focus is on capacities related to service delivery (i.e. technical capacities)	Impact and learning	• Both impact and learning are the foci, and the evaluation process itself produces lessons on process as well as impact.	• Story telling • MSC • Cust. satisfaction surveys • 360 degree audits • Empowerment evals. • Appreciative inquiry • Outcome mapping
Capacity monitoring	• Ongoing process for analyzing the extent to which an organization is developing, retaining and maintaining its capacities for immediate and long-term response. • Often uses an action plan to guide diligent capacity strengthening • Looks at outputs (from action plan) and outcomes primarily.	Learning	• Has a strong internal learning and organizational development focus, and looks for connections to impact in a more emergent manner (and sometimes doesn't make a formal attempt to link capacity and performance at all).	• Organizational self-assessments and action plans • Benchmarking against well-defined metrics • Appreciative inquiry

Action learning and capacity studies	• Better understand the relationship between capacity and performance by developing hypotheses and probing questions, and then testing them from capacities to impacts and vice versa, studying the process as intently as impacts produced.
	• Looks for performance breakthroughs and their links to differentiated organizational capacities.
	• Looks at the links between CD activities and capacities developed
		Learning	• This is the full CD learning agenda approach. Looks to answer questions from multiple angles, such as:
			  o What are the most critical capacities for high performance?
			  o To what extent are capacities dispersed throughout an organization?
			  o What are the linkages between individual, organizational, and system level capacities?
			  o What are the best methods for internal CD?
				• Action research
				• Use of evidence in relation to best practice standards
				• Outcome mapping

---

[27] Taken from a presentation on new trends in CD by Dr. Beryl Levinger, Distinguished Professor of Nonprofit Management at the Monterey Institute of International Studies. Additionally, this table was developed on the basis of a brainstorm conversation with Dr. Levinger. The assignment of specific methodologies to evaluation types does not necessarily reflect the views of the authors of any of the methodologies.

Capacity development is in motion, inching forwards, sideways, and sometimes backwards. Sometimes it has the wind behind it, sometimes it generates its own steam and sometimes it is stopped in its tracks by forces of nature, or man-made power structures. It lives within complex adaptive systems that ensure it will generally tend towards entropy and unpredictability. As such, the best way to penetrate the fog is by using methods that naturally thrive in the haze of complexity, and don't need to solve the capacity puzzle to prove that something worthwhile is happening. Open learning can coexist with complexity because it assumes complexity is the norm, and prefers to feel its way through the marshes (Bakewell and Garbutt, 2005) in order to maximize the possibility that CD interventions will be effective.

A learning approach to M&E of CD is intent on helping us get it right; not only getting it quantified and reported. It is about effectiveness but also about ownership. M&E primarily for donor accounting purposes puts in doubt where the ownership lies, as well as doubt as to the very purpose of CD. Contributing to endogenous development in motion assumes that local ownership of development is a desirable reality. M&E of CD primarily for learning takes a realistic approach to what can and should be measured and the reasons why. Donor accounting and transparency are also important, but should not be the primary focus of M&E of CD systems.

Ultimately, after a series of CD interventions (whether in a project program or other mode of activity) are complete, we believe that an evaluation of CD should ask the following questions:

- Did we thoughtfully analyse the related development history and stories before we started, including motivations, traumas, transformational moments, fears, wants, needs, etc.?
- Did we envision what success might look like from the perspective of the relevant stakeholders and design the interventions in such a way as to promote ownership of the processes?
- Were interventions selected with regard to how they built on processes already in motion, or at minimum, areas where there was a felt need or evident latent need for development—i.e. such that their exists an untapped, yet willing momentum for change?
- Did we debate progress, delays, insights, assumptions and activities frequently, and in a structured manner when possible?
- Did we change the project design often enough to keep it moving in a positive direction? Did we experiment enough?
- After all is said and done, did we get to a good place—did we do everything we could to promote the positive change we intended (broadly speaking)?
- What do we know now that we didn't know then and what would we do differently?
- Does everyone involved ask better questions as a result of this process?

The following map (figure 1) summarizes some of the principal dilemmas of M&E of CD described in this section:

Figure 1—Dilemmas with M&E of CD

4. Concluding thoughts

The discussion around M&E of capacity development reveals a core dilemma. Many development organizations consider CD 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 CD and why it is important (and spend large sums on money doing so), yet they rarely conceive of it, operationalize it, or measure it in a way that helps them learn and improve their approach.

Without a more strategic and operational understanding of CD, M&E efforts are unable to determine if CD efforts are leading us to understand development and development programming better. It is difficult to know if these efforts are informing our learning and adaptive management processes, directly or indirectly strengthening other development processes (and leading them towards impact), and developing or strengthening capacities that result in more system and organizational readiness and ability. Organizations are not systemically and systematically learning from and improving CD because they are unsure of what to look for.

So, what is at stake? In a recent evaluation of the World Bank's training processes, "It was found that most Bank-financed training resulted in individual participant learning, but improved the capacity of client institutions and organizations to achieve development objectives only about half the time (Independent Evaluation Group--IEG, 2008: xiii)". How many capacity development interventions continue to be carried out without serious reflection on what works and what does not?

There is an urgency to improve conception, design, implementation and learning in development programming in general, and capacity development's unique nature makes it the ideal "learning laboratory" for discovering key insights about development, and the links between capacity,

performance and impact. The CD that results from practicing "evaluative" thinking can change the way we look at development programming.

> "By providing a mechanism and process for clarifying values and goals, evaluation has an impact even before data are collected. Likewise, the process of designing an evaluation often raises questions that have an immediate impact on program implementation. Such effects can be quite pronounced, as when the process of clarifying the program's logic model or theory-of-action leads to changes in delivery well before any evaluative data are ever collected (Horton et al., 2003: vii)."

This urgency is reinforced at the very highest levels. Point 22 of the 2005 Paris Declaration on Aid Effectiveness highlights that "The capacity to plan, manage, implement, and account for results of policies and programmes, is critical for achieving development objectives — from analysis and dialogue through implementation, monitoring and evaluation."

The literature tells us further that M&E *must* play a central role in this process:

> "The M&E system should be a part of the strategic management of the programme from the outset. 'It must be part of an effort to embed and encourage an organisational culture oriented to performance, learning and self- reflection' (Morgan 1999:14). Organisations should be prepared to invest significant resources in this process (Roche 1999:260). (James, 2001: 4)"

These appear to be very rational suggestions, with little to argue against here. However, it is easy to overlook the fact that a fundamental shift in organizational culture is actually being demanded; that of "working smart", as opposed to simply "working hard"—learning as opposed to simply executing more projects.

> "The implications of this type of learning for an organization are thus less to do with knowledge management systems and processes, and more concerned with developing new tools for dialogue and holistic analysis, and attitudes and skills for working collaboratively (Pasteur, 2006: 22)."

In section 1 we stated the basic assumption that the end purpose of M&E is continual learning and improvement, manifested through change. This implies that improvement is possible, desirable, and necessary for finding the best way to complement and strengthen development processes and individual, organizational and system response-ability over time. Although the task is daunting, the current, largely ineffective status quo should not remain an option.

The following recommendations we believe are important for improved M&E for capacity development:

a) Incorporation of organizational learning approaches to M&E of CD
b) Large scale experimentation and action research
c) Use of Theory of Change Approaches for designing M&E of CD systems

4.1. Incorporation of organizational learning approaches to M&E of CD[28]

Although it is clear that the required shift to organizational learning for CD is more cultural than procedural, something needs to be done to start the process, and to test whether learning-based organizational cultures translate into more effective conception, design, implementation and evaluation of capacity development programming.

---

[28] We have already established throughout the importance of learning approaches to M&E for project/activity/intervention-level CD, and as such have not repeated that in this sub section.

The South Africa NGO CDRA's "homeweek" is an interesting organizational learning model, a dedicated monthly "week-long process of organisational connecting, strategising, action-learning, co-creating, managing, resource- allocating, peer-supervising, accounting, team building, record creating and practice developing (Soal, 2007: 3)."

> "As a consultancy our field staff are in a constant state of motion, in and out of the office, seldom keeping to ordinary office hours. Very early in CDRA's existence, consultants realised that they would seldom see each other, never build an organisational practice, if they did not commit to certain times in the year when they would all be in the office. Given the nature of consultancy - consultancy periods generally last from one to three working weeks - a monthly rhythm emerged as a sensible organising "structure-in-motion". Recognition was given that all organisational maintenance work – including learning, strategy, business and administration – should take no longer than a week. Thus the term "Homeweek" was coined. The form for learning that has evolved is contained within our "homeweek" – nine or ten weeks in the year, generally held towards the end of the month, in which all consultants are present in our offices in Cape Town and in which a variety of organisational and maintenance activities take place, including learning." (CDRA, 2001[29])

These activities are wide-ranging, including sharing of experiences by different members of the organisation, some of which in very creative ways; strategy and planning meetings; business meetings; and management and development of individual practice, such as coaching, mentoring and supervision. Reflecting the emerging nature of their work, "homeweek" is not predetermined. Instead, "[h]omeweek's spaces are shaped by the needs emerging out of our work in the field. The week becomes a melting pot, where the differentiated experience of each person, working alone, is shared with others, then actively forged into something else, something organisational (ibid: 6)."

Soal explains the transformative nature of this type of practice:

> "We have seen that organisations are best equipped to tackle problems of practice when they organize *themselves* to access the rich resource of experience that they already contain, and then translate into improved collective thinking, strategy and practice (ibid: 2)."

Another example of implementation of an integrated learning process is Action Aid's Accountability, Learning and Planning System (ALPS). ALPS is an attempt to move from a bureaucratic evaluation and planning system, to a learning based system that promotes more thoughtful analysis, debate, design, and adaptive management (David et al., 2006: 137). "ALPS was not designed to feed the insatiable desire for upward reporting (ibid: 140)."

> "...ALPS is based on sets of beliefs and principles around the rights of the poor to criticize and influence poverty eradication efforts. While turning the conventional "charitable" perspective on its head, ALPS offered little guidance and few rules. ALPS mainly opened up the space for creativity and provided a loose framework for country programmes to experiment with ways of improving development practice (ibid)."

ALPS is based on three core elements: principles; attitudes and behaviours; and organisational policies and processes. All these elements inform and guide the ways in which Action Aid approaches and carries out all of its activities, including with its partners, all of which aims to reinforce its accountability to the poor and excluded people with whom Action Aid and its partners work. These elements connect with all of Action Aid's organisational processes, including appraisals; strategies; strategic plans; annual plans and budgets; strategic reviews; peer reviews; organisational climate

---

[29]

http://www.cdra.org.za/nuggets/CDRAs%20homeweek%20%20by%20Sue%20Soal%20%20October%202001%20CDRA%20Nugget.htm accessed 16/9/08

reviews; annual participatory review and reflection; processes; annual; reports; internal governance annual review; external and internal audits; and open information policy.

By the fomenting of an "evaluative" culture, David notes four initial outcomes that are emerging from the ALPS process (particularly from the annual reflection process): "more learning, altered accountability, more transparency and enhanced organizational understanding of change and impact (ibid: 145)." These changes include improved decision making by virtue of getting relevant stakeholders involved in informing certain decisions.

While both of these systems are broader in scope than a pure CD focus, and both are full of lessons on the pain associated with cultural shifts (ibid: 143), they offer key insights on how learning approaches to CD can improve development programming. They both use iterative, continual stock taking approaches to determine the best way forward. Organizational priorities—defined emergently—drive the process, resulting in customized, self-led CD (allowing the organization time to accompany its own process). The core question is whether a mix of learning based approaches, widely adapted, could create large-scale system level improvements and impact. We cannot know what would happen unless more organizations[30] give it a try. To promote this, donors could include in the repertoire of CD interventions that they fund dedicated organizational learning time to process CD advances and improve decision-making. This could help reduce the over-reliance on consultants (while being able to more strategically utilize consultant time) by putting the onus of learning and change on the organizations themselves.

We recognise that national education systems fall largely within the public sector, and the challenges faced by educational institutions in achieving cultural shifts may be much greater than those encountered in organisations such as NGOs. However, if NGOs or other CD service providers can take on board a learning orientation to their M&E of CD, it may be possible for them to gradually introduce this also within the public sector organisations where they are forming partnerships and collaborations.

4.2. Large-scale experimentation and action research

The experimentation recommended at the organizational level also needs to be taken to a broader system level if M&E of CD is to have the wider impact we have argued it should.

At best, we imagine rather gradual progress towards an eventual tipping point when M & E of CD is conceived of and carried out in a radically different way. There is, however, a huge need for more ideas, experiments and lessons on how M&E for CD can be done differently, more effectively, and at different levels and scales. It is time to promote experimentation and research that leads to the production of evidence on what works and what doesn't, as well as innovative ways to measure differently and "reimagine" CD processes and their measurement:

> "Energy for good change exists in every context, but we must learn to construct approaches to detect the dynamics of specific context and to mobilise and nurture this energy productively through a process of dialogue. This means focusing on change and adaptive management in an approach rooted in endogenous strengths, needs, aspirations and expectations arising from specific contexts rather than seeing CD always from an exogenous, deficit perspective. We believe that a real sea-change may be achieved in how CD is understood and practiced, by: promoting empowering relationships; supporting rallying ideas; mobilising dynamic agents;

---

[30] We are cognizant that although the onus is on the "autonomous" organization to take the initiative in shifting towards higher impact processes, there are structural, often donor based limitations to this type of change. For example, a shift to a learning-based approach included a change in internal cost structures, and corresponding shifts in donor rules on funding use. A deeper discussion on how to promote this type of experimentation is necessary.

proactively framing and shaping the context for CD; enhancing grounding/enabling knowledge and skills through systemic learning processes (Taylor and Clarke, 2008: 4)."

Again, the subject is system level cultural change, and "[it's uncertain] whether development banks and donors themselves have the institutional capacity to cope with new paradigms of development cooperation based on trust and 'letting go' (Watson, 2006: viii)." There are obviously multiple reasons why donors tend not to favour more learning-based approaches over accounting approaches, in addition to issues of trust and letting go. Issues such as management culture, strong beliefs in data and figures, and emphasis on measurable accountability are also explanations. But although there are "few[31] examples of donors supporting informal monitoring using a systems thinking approach (ibid)", Watson points out:

> "There is persuasive evidence of the value and effectiveness - in contributing to organisational capacity building - of 'endogenous' M&E approaches that: are based upon participation through self- assessment of key players; encourage feedback, reflection and learning on the basis of experience; and promote internal and external dialogue between stakeholders (ibid)."

Wider scale testing and support for testing is clearly needed. Table 2 included the evaluation category "Action learning and capacity studies". Many of the problems with M&E of CD are conceptual, but many of the insights are field based. As such, action research is uniquely placed to help drive a broader learning agenda for improvement in M&E of CD and can help answer critical questions such as:

- What are the most critical capacities for high performance?
- What are the characteristics of organizations that are most likely to generate innovation and catalyze broad scale change?
- Do high capacity organizations working together create system level impact?
- What are the linkages between individual, organizational, and system level capacities?
- What is the most effective balance between process- and content-driven approaches for CD?
- Are large-scale CD programs the most effective way for increasing needed capacities at different levels, or would more targeted social innovation investment yield better results?

4.3. Use of Theory of Change (TOC) Approaches for designing M&E of CD systems

4.3.1. What can a Theory of Change offer?

Theory of Change (TOC) is a planning process for visualizing, constructing and substantiating the elements, conditions and interventions that are fundamental for achieving sustained impact. The principal advantage of TOC is that it encourages the design of a programmatic intervention model that includes the totality of the conditions necessary for the success of a program, project or overall organizational strategy (not only the areas that a particular project might typically take into account). Seeing the totality of conditions necessary for success, it becomes easier to evaluate whether the approaches and interventions that a particular project chooses to implement, in addition to assumptions of the roles and interventions of other actors, are sufficient to "cause" the success of our overall vision. This also reveals gaps in a particular CD project design—both at the level of conditions and interventions necessary for success—and makes it clearer to see what needs to be measured.[32] This may sound a recipe for complexity, but as the example included below (see 4.3.2) illustrates, the TOC is a powerful approach which helps us understand how change happens.

---

[31] Specifically referring to the ECDPM case studies.
[32] Paraphrased from an earlier writing of Alfredo Ortiz.

"We need good theories of social change for building the thinking of all involved in processes of development, as individuals, as communities, organisations, social movements and donors. The conventional division in the world today between policy-makers (and their theorising) and practitioners is deeply dysfunctional, leaving the former ungrounded and the latter unthinking [Reeler, 2007]."

The organization Keystone has an approach to measuring impact which echoes what many authors have said about the inherent limitations of using traditional M&E "donor accounting" approaches for measuring change of social change processes in complex adaptive systems.

"In most social change work, the problems are complex and not well defined. Solutions involve changing attitudes, relationships, capabilities, conditions and behaviours, and need to be worked out over time, with constituents, and often in collaboration with other organizations.

"Processes like these are best managed within the framework of a shared theory of change that guides planning, acting, reflecting and learning. Constituents first clarify a shared vision of success (or impact). Then they try to identify what change processes are already happening and how they work. Finally, they map pathways to outcomes – all the changes that they believe must take place in their context to achieve lasting success. These are observable changes, however small, in the conditions, behaviours, capabilities, attitudes, and relationships and conditions that are considered essential for long term success."[33]

Giving an example, the NGO Pact developed and tested a TOC framework for the Zambian HIV/AIDS Learning Initiative (ZHLI), a program focused on strengthening the operational, technical and financial capacity of Zambian NGOs, networks and ISOs leading multisectoral HIV/AIDS prevention and care activities.[34] In its TOC-based M&E framework the project intervention model was mapped, showing how program objectives and components, working with targeted actors and institutions, should plausibly effect change types (i.e. collaboration & alliances, culture, learning & innovation, policy, process & systems, strategy, structure and technology)—which should lead to broader HIV/AIDS service reach and quality. This included a look at how other actors relate to the program interventions, and a story based approach for collecting and validating program advances. The M&E approach was very experimental, but noted significant success as of the mid-term evaluation. .

 The TOC does of course need to be well grounded in a particular context, or it could appear somewhat abstract. For further elucidation we provide here also a hypothetical case of TOC as might be applied to EFA, drawing on IIEP's own vision, mission and strategy to provide us with a specific grounding.

[33] Taken from http://www.keystoneaccountability.org/node/34
[34] This report can be found at www.impactalliance.org. Additional program objectives include: "Fostering the development, testing, sharing and program integration of promising best practices and lessons learned for effective multisectoral response to HIV/AIDS; and, expanding collaboration and knowledge sharing by Zambian NGOs, networks and ISOs among all stakeholders leading multisectoral HIV/AIDS initiatives. BLOOM, E., KUMMER, B., REEVES, M. & BENNETT, C. (2006) Zambian HIV/AIDS Learning Initiative--Midterm Evaluation Report. Pact."

4.3.2. Hypothetical example of TOC use in EFA[35]

IIEPs mission is "to strengthen the capacity of countries to plan and manage their education systems"[36], and it does this by: "Training planners and managers in skills to analyze and plan, manage and implement, monitor and evaluate; Supporting institutions and improving administrative routines, organization, leadership skills; and, Fostering an enabling environment through policy forums, international co-operation and networking."[37]

On a basic, broad level the following diagram shows IIEPs vision for change and preconditions to change[38], without interventions included:

Figure 2—IIEP Change Vision
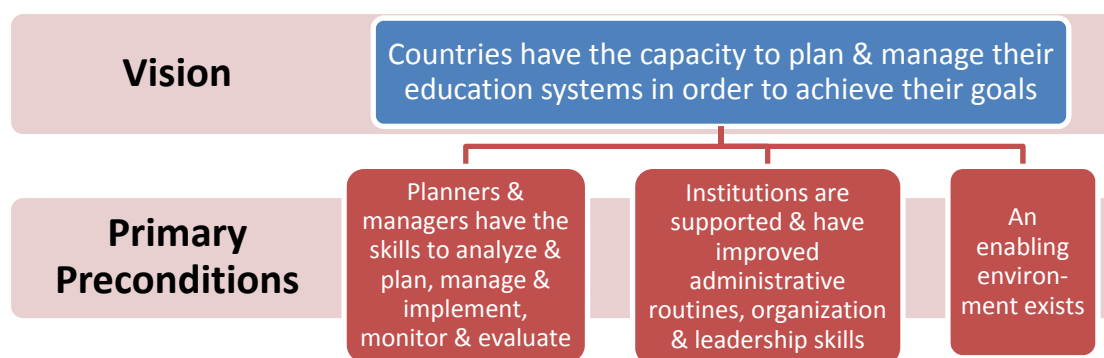
This basic model is essentially a partial theory of change (TOC) in that it includes a vision[39], it identifies the preconditions to success that must be in place before the vision can become a reality, and by adding interventions—which IIEP defines as capacity development in all of its core documents—IIEP will posit how it sees development happening, i.e. how it sees these conditions coming about, at least from the CD lens.

This is a shared TOC[40] in that IIEP is not the only organization that assumes the challenge of working towards country capacity to plan and manage education systems; so from the outset these preconditions include the potential results of IIEP *and* the results of other organizations and actors (both the governments themselves as well as other development and CD service providers) who are co-responsible for the achievement of this TOC. Broad, high level TOCs such as IIEPs assume the existence of complex adaptive systems by default, which adds an additional level of required M&E in order to know whether IIEPs CD is effectively filling its role—according to its own standards.

Within this broad TOC, IIEP has prioritized certain programmatic strategies and interventions. In its Medium-Term Plan 2008-2013, IIEP has prioritized scaling up training and organizational support for Member States; producing new knowledge in key areas for education policies, planning and management; and sharing knowledge in educational planning and management. This complements its ongoing training in skills to analyze and plan, manage and implement, monitor and evaluate;

---

[35] Although Education for All (EFA) emanates directly from UNESCO—it is in fact "the first Overarching Objective in UNESCO's Medium-Term Strategy for the period 2008-2013"—"as an integral part of UNESCO, the priorities of IIEP fit within those of the larger Organization [IIEP Medium-Term Plan 2008-2013]." As such, indirectly or directly, "IIEP's work will contribute to the achievement of these objectives…within its overall role of capacity development in the domain of educational planning and management. It will do so through its unique combination of training, research and technical assistance [ibid].

[36] Taken from IIEP website

[37] Ibid

[38] This is our paraphrasing of IIEPs original language for purposes of framing the M&E of CD issues.

[39] The existence of higher levels of vision—something that translates into quality of life for individuals, communities, etc.—isn't necessary for purposes of this example.

[40] As TOCs are "shared" by default.

supporting institutions and improving administrative routines, organization, leadership skills; and fostering an enabling environment through policy forums, international co-operation and networking.[41]

Utilizing this broad TOC framework we offer some observations and questions regarding the links between M&E of CD and EFA at the macro strategic, project and intervention, and partner levels.

Macro strategic level

M&E of CD Objective	Core questions
Help IIEP continually refine its strategic offering in relation to EFA.	• *Does our TOC accurately represent development complexities and realities regarding EFA?* ○ *Are our assumptions solid on how change happens between levels?* • *To what extent do IIEP's CD strategies respond to the needs of its shared TOC (and to what extent is IIEP's role adequately clear)?* • *Is the mix of what IIEP does, in conjunction with what others do, necessary and sufficient for the realization of the shared vision of success? (linkage to partner level)*
Description	
As part of a strategic management feedback loop, M&E should continually piece together evidence on whether the TOC assumptions and preconditions accurately reflect EFA realities and system level needs over time, and the extent to which the CD strategies IIEP chooses to pursue are the most effective ways for IIEP to catalyze this TOC. The answer to these questions should consider IIEPs mission and mandate, relative strengths and weaknesses, TOC "system" needs, and the roles of others in achieving change.	

Project and intervention levels

M&E of CD Objective	Core questions
Learning for improved relevance, effectiveness (impact) and efficiency of capacity strengthening interventions.	• *Do IIEP's CD program and project designs utilize the most effective interventions for achieving desired impact?* • *Does IIEP help build resilience, readiness and ability in the governments, organizations and individuals that it strengthens?* • *To what extent do IIEPs CD efforts take into account complex adaptive systems, including influences and relationships between levels of capacity (e.g. individual, organization, system)?* • *Does IIEP adequately utilize learning based approaches to M&E?*
Description	
This is in reference to the level of program and project interventions—i.e. the operationalising of IIEP's programmatic strategies. At this level, M&E of CD should help answer whether IIEPs selects the most adequate CD interventions to get to impact. In other words, do training, strengthening administrative routines, etc. effectively achieve desired impacts? Additionally, M&E should inform IIEP whether it builds adequate response-ability (i.e. standing capacities), including adaptive management abilities that help member states, support organizations' and individuals manage education priorities and relationships in complex, constantly adapting systems.	

---

[41] Taken from IIEP website

Partner level

M&E of CD Objective	Core questions
Learn from partner experiences and continually define the most ideal partners for responding to EFA.	• *What are we learning from our partners that can improve our information for decision making at an operational level. The policy level?* • *Do we need different, or additional partners for responding to areas of our TOC that are currently underserved?*
Description	
As part of a strategic partnership management, M&E of CD should aid IIEP in gathering lessons learned from implementing partners (govts, NGOs and others) to help inform its decision making at different levels. Additionally, analyzing M&E results at a cumulative level can help determine if IIEP's could benefit from additional and/or different partners to better respond to certain areas of the TOC (e.g. in the "enabling environment" precondition).	

As this example indicates, TOC thinking assumes that systems are complex and constantly adapting, that organizations are part of a broader system (and must learn to see themselves as such), and that thoughtful programming is emergent and learning based. This makes frameworks like TOC ideal for understanding how intangibles such as relational and adaptive management capacities need to be developed from an open system perspective informed by learning based M&E for CD.

4.4. Conclusion

Much of the literature states that the main uniqueness of M&E for CD (as opposed to for development in general) is that CD is more process oriented, therefore more in need of process type M&E. We are not convinced that this is much different from M&E of development in general. We believe the main uniqueness is the very learning nature of CD that, as a precondition to development, has the potential to help us better understand development, and more and more "get it right". Much is at stake, and M&E of CD can play a fundamental role in re-focusing towards learning and piecing together stories, thereby translating invisible fields into readable patterns that improve CD practice.

5. Acronyms

ALPS	Accountability Learning and Planning System
CAS	Complex Adaptive Systems
CB	Capacity Building
CD	Capacity Development
CDRA	Community Development Resource Association
CIDA	Canadian International Development Agency
ECDPM	European Centre for Development Policy Management
EFA	Education for All
IDRC	International Development Research Centre
IDS	Institute of Development Studies
IEG	Independent Evaluation Group
IIEP	International Institute for Educational Planning
INTRAC	International NGO Training and Research Centre
IT	Information Technology
M&E	Monitoring and Evaluation
MSC	Most Significant Change
NGO	Non-governmental organization
OECD	Organisation for Economic Cooperation and Development
RBM	Results Based Management
RFP	Requests for Proposals
SIDA	Swedish International Development Cooperation Agency
SP	Service Provider
TOC	Theory of Change
UNESCO	United Nations Educational, Scientific and Cultural Organization
US	United States
USAID	United States Agency for International Development
WBI	World Bank Institute
ZHLI	Zambian HIV/AIDS Learning Initiative

6. Bibliography

BAKEWELL, O. & GARBUTT, A. (2005) The use and abuse of the logical framework approach. Stockholm, Swedish International Development Cooperation Agency (SIDA).

BASER, H. & MORGAN, P. (2008) Capacity, Change and Performance: Study Report. IN ECDPM (Ed.) *Capacity, Change and Performance.* Maastricht, ECDPM.

BLOOM, E., KUMMER, B., REEVES, M. & BENNETT, C. (2006) Zambian HIV/AIDS Learning Initiative--Midterm Evaluation Report. Pact.

BUCKINGHAM, M. & COFFMAN, C. (1999) *First, Break all the Rules: What the World's Greatest Managers do Differently*, Simon and Schuster.

CAILLODS, F. & DE GRAUWE, A. (2006) Capacity Development in Educational Planning and Management: IIEP's Experience. *International Working Group on Education.* Rome, IIEP UNESCO.

COLLINS, J. (2005) *Good to Great and the Social Sectors*, Jim Collins.

CRUTCHFIELD, L. R. & GRANT, H. M. (2008) *Forces for Good: The Six Practices of High-Impact Nonprofits*, Jossey-Bass.

DAVID, R., MANCINI, A. & GUIJT, I. (2006) Bringing Systems into Line With Values: The Practice of the Accountability, Learning and Planning System (ALPS). IN EYBEN, R. (Ed.) *Relationships for aid.* London, Earthscan.

DAVIES, R. & DART, J. (2005) The 'Most Significant Change' (MSC) Technique: A Guide to Its Use.

DLAMINI, N. (2007) Stories and M&E. *ONTRAC.* INTRAC.

EARL, S., CARDEN, F. & SMUTLYLO, T. (2001) *Outcome mapping: building learning and reflection into development programs,* Ottawa, IDRC.

EARLE, L. (2002) Lost in the Matrix: The Logframe and the Local Picture. INTRAC.

EYBEN, R., KIDDER, T., ROWLANDS, J. & BRONSTEIN, A. (2008) Thinking about change for development practice: A case study for Oxfam GB. *Development in Practice,* 18**,** 12.

HAILEY, J., JAMES, R. & WRIGLEY, R. (2005) Rising to the Challenges: Assessing the Impacts of Organizational Capacity Building. IN INTRAC (Ed.) *Praxis Papers.*

HORTON, D., ALEXAKI, A., BENNETT-LARTEY, S., BRICE, K. N., CAMPILAN, D., CARDEN, F., SILVA, J. D. S., DUONG, L. T., KHADAR, I., BOZA, A. M., MUNIRUZZAMAN, I. K., JOCELYN PEREZ, CHANG, M. S., VERNOOY, R. & WATTS, J. (2003) *Evaluating Capacity Development: Experiences from Research and Development Organizations around the World*, ISNAR, CTA, IDRC.

INDEPENDENT EVALUATION GROUP--IEG (2008) Using Training to Build Capacity for Development: An Evaluation of the World Bank's Project-Based and WBI Training. The International Bank for Reconstruction and Development / The World Bank.

JAMES, R. (2001) Practical Guidelines from the Monitoring and Evaluation of Capacity Building: Experiences from Africa. IN INTRAC (Ed.) *Occasional Paper Series.* Oxford, INTRAC.

KAPLAN, A. (1999) The Developing of Capacity. CDRA.

LAFOND, A. & BROWN, L. (2003) A Guide to Monitoring and Evaluation of Capacity-Building Interventions in the Health Sector in Developing Countries. *MEASURE Evaluation Manual Series.* Carolina Population Center, UNC Chapel Hill.

LESSIK, A. & MICHENER, V. (2000) Measuring Institutional Capacity. IN (CDIE), U. C. F. D. I. A. E. (Ed.) *Recent Practices in Monitoring and Evaluation TIPS.* USAID.

LUSTHAUS, C., ADRIEN, M.-H. & PERSTINGER, M. (1999) Capacity Development: Definitions, Issues and Implications for Planning, Monitoring and Evaluation. *Universalia Occasional Paper.*

MEBRAHTU, E., PRATT, B. & LÖNNQVIST, L. (2007) *Rethinking Monitoring and Evaluation*, INTRAC.

MORGAN, P. (1997) The Design and Use of Capacity Development Indicators. Gatineau, CIDA.

MORGAN, P. (2006) The Concept of Capacity. *Capacity, Change and Performance.* Maastricht, ECDPM.

OECD (2006) *The challenge of capacity development working towards good practice,* Paris, OECD.

PASTEUR, K. (2006) Learning for Development. IN EYBEN, R. (Ed.) *Relationships for aid.* London, Earthscan.

SEN, A. (1999) *Development as Freedom*, Anchor Books.

SOAL, S. (2007) Towards "better evaluation"—an account of one internal practice. *Australasian Evaluation Society.* Melbourne, CDRA.

TAYLOR, P. (2008) Higher education curriculum for human and social development. *GUNI--Higher Education for Human and Social Development: new roles and emerging challenges (Palgrave),* 2008**,** 12.

TAYLOR, P. & CLARKE, P. (2008) *Capacity for a Change,* Brighton, IDS.

VANSANT, J. (2000) A Composite Framework for Assessing the Capacity of Development Organizations.

WATSON, D. (2006) Monitoring and evaluation of capacity and capacity development. IN ECDPM (Ed.) *Capacity, Change and Performance.* Maastricht, ECDPM.

WRIGLEY, R. (2006) Learning from Capacity Building Practice: Adapting the 'Most Significant Change' (MSC) Approach to Evaluate Capacity Building Provision by CABUNGO in Malawi. *Praxis Papers.* INTRAC.