



United Nations
Educational, Scientific and
Cultural Organization

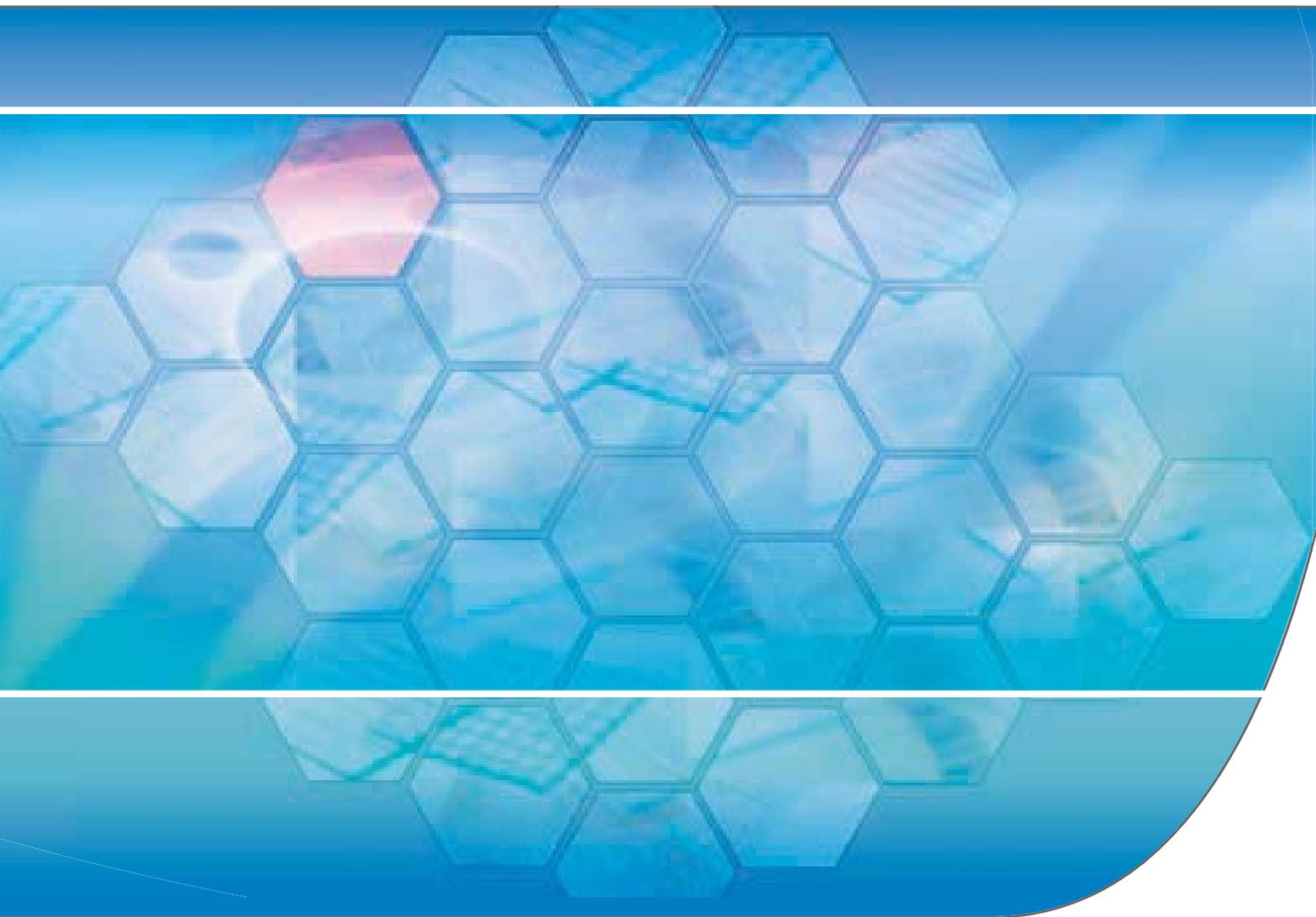


International Institute
for Educational Planning

Learning purposefully in capacity development

Why, what and when to measure?

An opinion paper prepared for IIEP, by
Alfredo Ortiz and Peter Taylor,
Institute of Development Studies, UK



Learning purposefully in capacity development
Why, what and when to measure?

Learning purposefully in capacity development

Why, what and when to measure?

An opinion paper prepared for IIEP

by Alfredo Ortiz and Peter Taylor
Institute of Development Studies, UK



United Nations
Educational, Scientific and
Cultural Organization



**International Institute
for Educational Planning**

The views and opinions expressed in this book are those of the authors and do not necessarily represent the views of UNESCO or IIEP. The designations employed and the presentation of material throughout this review do not imply the expression of any opinion whatsoever on the part of UNESCO or IIEP concerning the legal status of any country, territory, city or area or its authorities, or concerning its frontiers or boundaries.

The publication costs of this study have been covered through a grant-in-aid offered by UNESCO and by voluntary contributions made by several Member States of UNESCO, the list of which will be found at the end of the volume.

The author would like to thank Lyndsay Bird, Peter Buckland, Ahmed Ferej, Dorian Gay, Colin Kaiser, Khalil Mahshi and Maurice Robson for sharing their thoughts and experiences related to improving capacity development activities in fragile situations. Thanks also to Anton De Grauwe for helpful comments and editing of the publication.

Published by:
International Institute for Educational Planning
7-9 rue Eugène Delacroix, 75116 Paris, France
info@iiep.unesco.org
www.iiep.unesco.org

Cover design: IIEP
Typesetting: Linéale Production
Printed in IIEP's printshop
iiep/web/doc/2009/11
© UNESCO 2009

Contents

List of abbreviations	6
List of tables	7
List of figures	7
List of boxes	7
Foreword	8
Summary	9
Résumé	10
Introduction	
What should monitoring and evaluation tell us about capacity development?	11
1. Capacity development means and ends: what are we measuring and when should we measure it?	16
In search of performance and impact	16
Standing capacity	21
2. What can we learn from the dilemmas of the monitoring and evaluation of capacity development?	25
Development being a process already in motion	25
Linear versus complex adaptive systems thinking, programming and measurement	26
Attribution	29
Donor accounting focus versus open learning approaches	30
3. Concluding thoughts	37
Incorporation of organizational learning approaches to the monitoring and evaluation of capacity development	38
Large-scale experimentation and action research	40
Use of theory of change approaches for designing the monitoring and evaluation of capacity development systems	41
Conclusion	46
Bibliography	47

List of abbreviations

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
DOSA	Discussion Oriented Self Assessment
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
MDGs	Millennium Development Goals
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
UNDP	United Nations Development Programme
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

List of tables

Table 1.1	CD interventions, relations and measures	18
Table 2.1	Selected evaluation types, and different focuses and uses for capacity development	34

List of figures

Figure 2.1	'Feeling the way to the goal'	28
Figure 2.2	Dilemmas with the monitoring and evaluation of capacity development	36
Figure 3.1	Simplified TOC diagram (adapted from ActKnowledge, 2009)	42
Figure 3.2	IIEP change vision	44

List of boxes

Box 1.	A clarifying note on TOC diagram structure and non-linearity	43
--------	--	----

Foreword

Capacity development is a fundamental part of the mandates of many international organizations. Much of their work aims to strengthen national capacities through training, technical advice, exchange of experiences, research, and policy advice. Yet there is considerable dissatisfaction within the international community regarding the impact of many such interventions. The activities have usually strengthened the skills of individuals, but have not always succeeded in improving the effectiveness of the ministries and other organizations where those individuals are working. These shortcomings demand investigation in order to strengthen capacity development policies and strategies.

In this context, UNESCO received funds from the Norwegian Ministries of Education and Foreign Affairs to focus on 'capacity development for achieving the Education for All goals'. The objective was to identify appropriate strategies for UNESCO and others. Within UNESCO, IIEP has coordinated this work. A wide range of activities was undertaken, including detailed case studies on three countries (Benin, Ethiopia and Vietnam), a series of thematic studies and literature reviews, and consultations with experts. The focus has been on educational planning and management as stronger capacities in these areas should lead to important improvements in the education system as a whole.

IIEP's work has led to the identification of some main principles:

- The type of capacity development being considered here only works in a sustainable manner when there is national leadership and ownership, and when international efforts match national priorities and strategies.
- Strategies need attention at several levels: the capacities of the individual, the effectiveness of the organization (for example the ministry of education), the norms and practices which rule public management as a whole, and the political, social and economic contexts.
- Any intervention must recognize the intrinsic values of ownership and participation. When it aims only to identify partners' weaknesses or to strengthen the positions of those already powerful, the deepest sense of capacity development is lost.

The series *Rethinking capacity development* has been prepared within this framework.

Mark Bray
Director
UNESCO/International Institute for Educational Planning

Summary

Many capacity development (CD) processes aim at long-term sustainable change, which depends on seeing many smaller changes in what are often intangible fields (rules, incentives, behaviours, power, coordination etc.). Yet, monitoring and evaluation (M&E) processes of CD tend to focus on short-term outputs and clearly visible changes.

This paper offers some ideas on how to deal with this paradox by delving into what CD elements M&E can and should be able to measure. This, in turn, depends on whether CD is considered a means or an end.

The paper explores predominant ways of thinking about M&E, and discusses whether these are able to grasp the unique nature of M&E of CD. M&E should be able to measure how CD contributes to wider development processes and to sustainable capacity, in addition to measuring the quality of the CD process itself. Breaking down the concept of CD may be useful in order to observe CD ends. It may also be useful to gear CD more towards nurturing long term, even unintended outcomes. The notion of standing capacity is useful in order to measure capacity beyond pre-programmed, immediate performance.

The authors further draw a number of lessons linked to current evaluation dilemmas: evaluating the fit between CD efforts and ongoing development processes; taking into consideration that CD is not always a linear process; and understanding that the difficulty of attributing changes to specific CD efforts is an important clue as to how CD works.

A broader approach to M&E of CD may be necessary in order to integrate these lessons: Open systems thinking and learning approaches through the use of storytelling may prove to be more strategic and efficient than the instrumental approaches often used by donors. A hypothetical example of Theory of Change use in EFA equally demonstrates the need for new thinking in order to improve M&E, and consequently, CD.

Résumé

De nombreux processus de renforcement des capacités (RC) ciblent des changements durables sur le long terme, ce qui implique l'observation de nombreuses transformations sur une petite échelle dans des domaines souvent intangibles (règles, incitations, comportements, pouvoir, coordination, etc.). Néanmoins, les processus d'évaluation du RC ont tendance à se focaliser sur des résultats axés sur le court terme et des transformations clairement visibles.

Cet article propose quelques idées pour traiter ce paradoxe en creusant les éléments de RC que le *monitoring et l'évaluation* (M&E) peuvent et devraient pouvoir mesurer. Ceci, à son tour, dépend de ce que l'on pense être le résultat final de RC, et dans quelle mesure le RC peut être considéré un moyen ou une fin.

Le document examine si les approches prédominantes dans le M&E sont capables de saisir la nature unique du M&E du RC. Selon les auteurs, le M&E doit être capable de mesurer la manière dont le RC contribue aux processus plus larges de développement et au développement de capacités durables, ainsi que la qualité même du processus de RC. Décomposer le concept de RC peut s'avérer utile afin d'observer les finalités qui lui sont attachées. Il est aussi utile d'orienter le RC davantage vers l'encouragement des résultats sur le long terme, voire des résultats inattendus. La notion de capacité mobilisable est utile afin de mesurer des capacités au-delà de la performance immédiate et préprogrammée.

Les auteurs identifient ensuite un certain nombre de leçons liées aux dilemmes contemporains de l'évaluation : évaluer la manière dont un effort de RC correspond aux processus de développement déjà amorcés ; prendre en compte le processus souvent non linéaire de RC; et comprendre que les difficultés liées à l'attribution des transformations aux efforts de RC constituent un indice important quant au fonctionnement du RC.

Une approche plus large du M&E sera peut être nécessaire afin d'intégrer ces leçons: les approches axées sur l'ouverture des systèmes et l'apprentissage à travers la narration peuvent être plus stratégiques et plus efficaces que les approches instrumentales souvent utilisées par les donateurs. Un exemple hypothétique de l'utilisation de la théorie de changement en EPT démontre également le besoin d'une réflexion nouvelle afin d'améliorer le M&E et, par conséquent, le renforcement des capacités.

Introduction: What should monitoring and evaluation tell us about capacity development?

The debates around the monitoring and evaluation (M&E) of capacity development (CD) are linked inextricably to the larger debates around CD and M&E individually. To monitor and evaluate CD presumes having a strong idea as to what its ‘means’ and ‘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 and more or less time restricted, and so on. But before the monitoring and evaluation of CD is delved into, it is important to define what is meant by ‘capacity’.

Capacity has been defined from the very broad, for example: “[C]apacity’ is understood as the ability of people, organizations and society as a whole to manage their affairs successfully” (OECD, 2006: 12) to the more narrow, for example: “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, for example: “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). It has also been defined with minimalist simplicity, for example: “Capacity is [the] potential to perform” (Horton, Alexaki, Bennett-Lartey, Brice and Campilan *et al.*, 2003: 18) It has been divided into ‘hard’ capacities such as “infrastructure, technology, [and] finances” (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 those which might appear to be more ‘tangible’ – that is, the systems and processes mentioned above, and ‘intangible’ – that is, capacities that 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). These capabilities, in addition to the capability “to carry out technical, service delivery and logistical tasks,” form the European Centre for Development Policy Management (ECDPM) capabilities model. (See also the insightful table on adaptive capacities in Horton *et al.*, 2003: 28-29.) 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” (www.merriam-webster.com/dictionary/capacity).

Notions of capacity also relate to broader issues of human development. Although hotly debated over the decades, development has been equated by many with global economic growth, which would result in all people achieving economic parity with those living in the ‘developed’ nations. Over time, ‘human development’ has, however, acquired a more complex meaning. 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, n.d.).

The UNDP statement, echoing Sen (1999), 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 (UNDP, n.d.).

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 and the UNESCO Decade for Sustainable Development. These frameworks do not guarantee 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* among peoples to alleviate conflict and violence, *equality based on gender, ethnicity and religious belief*, and even the striving for *‘happiness’*. All play a part in determining global development pathways, yet all are influenced to a greater or lesser extent by *capacity* (Taylor, 2008).

The many angles of capacity demonstrate the richness and yet daunting nature of the subject. Monitoring and evaluation are fundamentally about measurement, which we look to in order to help decipher this complex puzzle of CD. What needs to be defined in order to improve CD design, implementation, learning, performance and impact is when, what and how to measure. What can or should M&E contribute to ensure that CD is understood better and is 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 (which are answered briefly in this section and in more detail throughout the paper, while the remaining sections of the paper are not structured on the basis of the assumptions) about CD in the broad field of education, and indeed in other sectors:

- a) Monitoring and evaluation of CD should help us understand development and development programming better, and inform our learning and adaptive management processes.
- b) Monitoring and evaluation of CD should show whether CD processes strengthen other development processes directly or indirectly and guide them towards making an impact in the most thoughtful way possible.
- c) Monitoring and evaluation of CD should aid in demonstrating whether CD strengthening processes are supporting the development of capacities that result in more system and organizational readiness and ability – or ‘response-ability’ (Fowler, 2000: 8), including the development of reserve and emergency capacity.

In this paper, we distinguish between *monitoring* and *evaluation*. Monitoring implies an ongoing measurement process and evaluation a periodic measurement process. Throughout the paper we will use the term ‘measurement’ as a broad reference to M&E, in addition 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. We consider each of the three key assumptions in turn:

“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, Anderson, and Murphy, 1995: 15).

a) Monitoring and evaluation of CD should help us understand development and development programming better, and inform our learning and adaptive management processes.

“When the final word is written about capacity building ... what will the record show? For many practitioners, from whatever role – consultant or technical assistance provider, grant maker or researcher – the most compelling test will be whether organizations and the sector as a whole have become stronger and more effective in their efforts. Will we have improved the quality of life in the communities where we work? Will we have contributed to society’s willingness to embrace systems change and sustainable solutions ...?” (Linnell, 2003: 5).

Observing thoughtfully how capacity develops – increases, recedes, take leaps or evolves – over time is an exercise in seeing whether or not some of the most important preconditions for development are present. 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 correlation between capacity and development, and M&E should help document and process those lessons. It should

tell us about what works in development and 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 or ecosystem is helping or hindering capacity development efforts. The key elements to be measured here are *learning for development and adaptive management*.

b) Monitoring and evaluation of CD should show whether CD processes strengthen other development processes directly or indirectly and guide them towards making an impact in the most thoughtful way possible.

Whether capacity is a means to or an end of development, or both – and this paper strongly supports the assertion that it is both – over time its ultimate purpose is to enhance or strengthen the ability to achieve other development ends. (See the section on ‘standing capacity’ for a more detailed discussion on capacity as an end of development.) It attempts this in some cases by trying to improve, replicate or scale up other primary development programmatic interventions, for example improving service delivery in health or education via training. In other cases, it does not aim to strengthen specific primary development processes themselves, but to enhance *the ability* of different actors to strengthen these processes in a more macro sense (this is explored further in the next question). Better developed capacity should eventually have a positive effect on other development processes, and M&E processes should show the extent to which CD interventions contributed to enabling these development processes.

Additionally, M&E processes should measure the quality of the capacity strengthening process itself, including the extent to which CD practitioners (regardless of whether they are local or external) design, facilitate and improve their processes taking into account ongoing development, its participants, its history and its inherent complexity. 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) Monitoring and evaluation of CD should aid in demonstrating whether CD strengthening processes are supporting the development of capacities that result in more system and organizational readiness and ability – ‘response-ability’, including the development of reserve and emergency capacity.

Individuals and organizations need capacity well beyond what they use on specific projects or activities each day. A predominant focus on directly linking capacity development to immediate technical performance might lead to a lack of preparedness for atypical situations – which occur throughout everyday life. In addition to improving today’s technical performance, strengthening processes should help develop broader ‘standing capacities’, including key intangible qualities such as relationship leverage, programme design capabilities, innovative culture, autonomous self-motivation and agile, adaptive management (Kaplan, 1999; Morgan, 2006). ‘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 *Chapter 1*. 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. The key elements to be measured here are *resilience, readiness and ability in responding to development challenges over time*.

In this paper, we take the position that the monitoring and evaluation of CD should be more about understanding what is *worth* measuring than about what *can* be measured simply by applying comprehensive M&E tools and indicators. We start with the assumption that the end purpose of M&E is continual learning and iterative adjustment for positive change. This implies that improvement is possible, desirable, and necessary for finding better ways to complement and strengthen development processes and individual, organizational and system response-ability over time. This also implies that learning approaches to both CD and M&E are needed to support this change. We assume that capacity is for *performance* – that is, understood as an organization doing its work effectively – but we recognize 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 effectively in the long term.

The remainder of this paper will explore in more detail the meaning of ‘capacity’ and its links to organizational effectiveness and impact, in order to clarify what is useful to measure, when and why. The following chapter will define capacity development ends in more detail in order to clarify what is being measured under the rubric ‘monitoring and evaluation 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. *Chapter 2* focuses on what can be learned from key elements and dilemmas in the debate on M&E of CD. *Chapter 3* 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 it 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 have implications for approaches to measurement and, we hope, will generate further interest in continuing this field of discussion.

1 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 in which the projects were developed, and educating and training high-level specialists in developing countries. Yet the lack of capacity to design, plan and implement policies and reforms is arguably one of the most serious obstacles to the implementation of EFA in developing countries today (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 the thinking and practise 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 – particularly at the individual level. However, as in other sectors, education policy, systems, processes and structures need to be examined to reveal the extent to which the monitoring and evaluation of CD supports the broader learning agendas that it aims to support. An examination of the relation of monitoring and evaluation of CD with 'performance' and 'impact' 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*).

In search of performance and impact

The European Centre for Development Policy Management (ECDPM) capabilities model noted earlier balances out its broad mix of soft, hard and intangible capabilities (that is, 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 achieve their missions directly. 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* – that is, an organization effectively doing what its mission or mandate says it does – and ultimately supporting emergent, positive change. If this is true, then for monitoring and evaluation of CD, the two relevant questions regarding 'capacity' are:

- a) Is capacity as a broad concept useful in determining what to measure (on the path to performance and positive, emergent change)? Is it helpful to break down the broad concept of capacity in order to measure specific dimensions of capacity?
- b) To what extent should capacity interventions be geared towards *immediate* performance?

We now consider these questions in some detail.

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, and so on (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 of 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, as Morgan points out, does 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’ (Horton *et al.*, 2003: 18) moves us one step further in that it describes capacity as a latent state, the potential energy of which 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 end and know that its accomplishment is a proxy for the pre-existence of capacity (LaFond and Brown, 2003: 8)? If capacity is an end in its own right – about expanding choices through the development of human capabilities, that is, “the range of things that people can do or be in life” (UNDP, n.d.) – then how should it be measured and should measurement attempt to link 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 of the impact?

Baser and Morgan’s 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 and system), *types* of capacity (competencies, capabilities, assets and relationships) and *purpose* (enabling human systems to create value). Thinking about measuring the development of individual competencies or the strengthening of organizational capabilities seems less daunting than thinking about measuring ‘capacity’ overall. The more micro and specific we go – for example, the capacity to implement procurement procedures – the more straightforward the task of M&E could potentially become.

Table 1.1 illustrates different types of breakdowns by capacity development purpose and shows how different types of capacity development interventions have different M&E implications. This table is not intended to be comprehensive; it aims only to draw the connection between different areas of capacity and M&E.

Table 1.1 CD interventions, relations and measures

Purpose of CD intervention	Intervention examples	Complexity factor ⁴ on achieving impact	Risks and relations to other capacities ²	Measures
Broadly strengthen organizational capacities and increase organizational learning	<ul style="list-style-type: none"> 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 DOSA³ self assessment tool and process, which analyses human resource management, financial resource management, service delivery, external relations, strategic management & organizational learning (VanSant, 2000). 	<ul style="list-style-type: none"> High in general because of the tenuous links between broad organizational capacities and development ends; extremely high if the organization doesn't have internal capacities to address the issues identified as key weaknesses⁴ (e.g. a specialized content-driven area such as financial management, marketing or a mission-related technical area). 	<ul style="list-style-type: none"> 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 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 analysed. A key benefit is the learning orientation that helps personnel from different horizontal and vertical layers see themselves as part of a broader organization. 	<ul style="list-style-type: none"> 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 reassesses itself periodically.
Strengthen organizational technical capacities	<ul style="list-style-type: none"> Training of individuals in particular techniques. For example training a teacher in a new lesson planning process or a programme technician in business planning for microenterprises. 	<ul style="list-style-type: none"> 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 functioning). 	<ul style="list-style-type: none"> Organizations that overly focus on technical capacities without strengthening the organizational systems and processes that support technical execution.⁵ On a practical level, risks are high because of potential transfer problems between the individual and the organization (assuming the training was sound). 	<ul style="list-style-type: none"> 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).

Purpose of CD intervention	Intervention examples	Complexity factor ¹ on achieving impact	Risks and relations to other capacities ²	Measures
Strengthen financial stability	<ul style="list-style-type: none"> • Training or coaching in fundraising, marketing, strategic communications and negotiation skills (e.g. with ministries and other funders), for example. 	<ul style="list-style-type: none"> • High. Multiple external factors – in addition to improved fundraising, marketing or communications techniques – affect whether an organization will be successful in improving financial stability. 	<ul style="list-style-type: none"> • 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 offer high-value development contributions 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. 	<ul style="list-style-type: none"> • Increased funding, diversification of funders, core (unrestricted) funding and alignment of funding with organizational core competencies and priorities (i.e. increased autonomy).
Strengthen a specific internal process (e.g. procurement)	<ul style="list-style-type: none"> • Process to map the purchasing cycle and document procedures for different types of purchases (e.g. products and services) and different threshold/risk levels. 	<ul style="list-style-type: none"> • Medium to low – mostly internal organizational issues that could affect successful implementation. 	<ul style="list-style-type: none"> • Most purchasing processes don't analyse cost-benefit in a macro sense (as long as it's in 'the budget' the legitimacy of the purchase is accepted). 	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • Participatory processes that link actors within a system and build or make visible areas of common ground. • Processes that help managers systemically analyse decisions, improve information for decision-making and strengthen scenario planning. 	<ul style="list-style-type: none"> • High for both, multiple factors outside the control of the organizations. 	<ul style="list-style-type: none"> • Highly dependent on individual talents and motivations, many of which are difficult to cultivate if they don't already exist (Buckingham and Coffman, 1999). 	<ul style="list-style-type: none"> • Very difficult to measure conventionally and outcome measures (e.g. participation in networks and more informed decisions) are potentially meaningless without an accompanying story or narrative.

1. 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.
2. By risks and relations we mean to identify the relationship between this particular area of capacity and others, and the risks posed by approaches that are too narrow or broad.
3. 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).
4. 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.
5. 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.

A breakdown of 'capacity' is helpful because the term then becomes easier to understand and is more relevant at an intervention level, leading logically to the types of measures that will indicate effectiveness. At this level it becomes actionable – the 'capacity for what' becomes clearer and more intentional. Taking a micro-level standpoint runs 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 analysing capacity development complexities, but it must be grounded in the reality of actual interventions within a particular context.

b) To what extent should capacity interventions be geared towards immediate performance?

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 programmes usually view institutional capacity as a means to achieve higher level programme 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, based on Lusthaus, Anderson, and Murphy (1995) and Lusthaus *et al.* (2002).

Performance as a concept might unwittingly encourage thinking of capacity in terms of short-term or limited technical output. 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). Horton *et al.* express it as "an organization's performance ... in terms of four key indicators: effectiveness, efficiency, relevance and financial sustainability" (2003: 20).

The difference between these views of performance is important. In the Baser and 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 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 it may or may not yield. It can potentially get 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 seen only 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 or 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 and programmable. 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 to strengthen an organization’s ability to contribute to positive change. 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 will only set us up for failure when it comes to the time for measurement.

Standing capacity

The idea of ‘standing capacity’, as 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 a lack of standing capacity quickly reveals woefully inadequate responses that affect thousands of lives. Yet 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 the potential disruption of educational provision for a whole generation of school-aged children?

We divide standing capacity into two areas: (1) basic functionalities; and (2) organizational talents, both of which are important for good performance. 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 knows how to do; and talent – what one does, thinks or feels consistently (well) and can put to productive use. Basic functionalities are the minimal systems, 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. Organizations that are under-resourced, overcommitted, poorly managed, donor driven or misplaced (for example, an education institution that finds itself manufacturing local products, which takes 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 respond to everyday needs. Some sports provide a good metaphor for reasonable excess capacity. In professional basketball, for example, although only five players play at a time, and very few teams rotate in more than ten 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 individuals’ lack of capacity 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 decentralizing policies, which encourages 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.

The second level of standing capacity is what we refer to as ‘organizational talents’. This represents an organization’s ability to summon, draw upon or leverage a unique combination of capabilities, resources, synergies, intuitions and relationships, allowing 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. 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. Organizational talents are about highly effective, 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 prepare only for their limited results and immediate programming, then they are not preparing systemically. Monitoring and evaluation of CD needs to measure standing capacity in order to better understand the links between capacity and performance. 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 project management or procurement systems and other harder capacities.

Standing capacity is important at a higher level as well. On a macro level, if development is expected to result in a broadening of freedoms (Sen, 1999), then the capacity of institutional systems – with related actors, incentives, norms, processes and broader environmental issues – 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 broader systems are ready and able, both for running individual clinics more efficiently today as well as 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 ten years (for example, a public works project that builds levees that are 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 (that is, immediate performance obsession), are strengthening system-level standing capacity.

Readiness and ability are preconditions to guaranteeing freedoms, even when those freedoms are not immediately threatened nor tied to an immediately identifiable sectoral outcome. There is an assumption that capacities will eventually be put to use and 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. 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. 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 capacity, the concept used here includes the capacities of system-level actors and formal governance processes, as well as the capacities of individual actors to 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; Bloom and Dees, 2007; 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 the 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, that is, 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.”

As discussed earlier, the monitoring and evaluation of CD is useful to the extent that it helps change related behaviour – that is, improves CD design, implementation and learning – or improves

accountability to the 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.

[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).

2 What can we learn from the dilemmas of the monitoring and evaluation of capacity development?

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 back to donors (that is, legislatures and taxpayers in some cases) as to the efficacy of the use of funds. Whether it involves the 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 the donor intended*.

The problem lies with ‘as the donor intended’. Donors and sub-donors tend to create tremendous pressure for those whom they support to plan and measure much more than they could possibly – and intelligently – predict. The consequences of this pressure are manifested in a huge effort by implementing actors to show results in order to comply with sanctioned programming mandates (for example, 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 having a plan, resources 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).

An examination of some of these dilemmas shows how they affect the way we understand, plan, execute and measure CD. Even the way we speak (for example, “... affect the way we understand, plan, execute and measure CD”) is indicative of how prevalent externally driven CD processes are in many contexts. Four dilemmas are particularly relevant for the monitoring and evaluation of CD:

“[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).

- 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.

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 the arrival of any project, donor, programme or initiative, and not to recognize this as 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 executed in any event, yet 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; that is, 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; that is, positive development initiatives and processes already in motion.
- 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 in 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.

The monitoring and evaluation of CD should test periodically whether the selected interventions are likely to take hold based on their fit or relevance to ongoing development processes. There is a need to discover the extent to which CD processes are making an honest effort at thoughtfully 'reading' the situation (Kaplan, Morgan and multiple authors cite the importance of 'reading' the context before intervening) before acting. 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 and so on), before talking further about design. Furthermore, she insisted that we structure the workshop component of the processes in such a way as to minimize the interruption of work cycles. She was happy about the process, but felt that past external processes had been *applied to* her organization in a predetermined manner, not building on existing processes and momentum and not respecting the way the organization worked and learned best (intermittent workshops versus intense 'lockdowns'). This requires spending time and resources to study the context before intervening, as well as 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 these approaches 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.

"There is a need to observe and understand the change processes that already exist in a living social system. If we can do this before we rush into doing our needs analyses and crafting projects to meet these needs, we may choose how to respond more respectfully to the realities of existing change processes rather than impose external or blind prescriptions based on assumed conditions for change" (Reeler, 2007: 2)

Linear versus complex adaptive systems thinking, programming and measurement

Apart from being unable to predetermine precise long-term outcomes without knowing where one is starting from, it is in any case inadvisable to attempt to predict these outcomes *precisely* because multiple, unpredictable factors are often at play, influencing 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 (see the column 'Complexity factor on achieving impact in *Table 1.1*). One of the basic premises of complexity theory for CD is that *the direction in which 'development' is going has 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. Organized 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: 203-204).

"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).

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

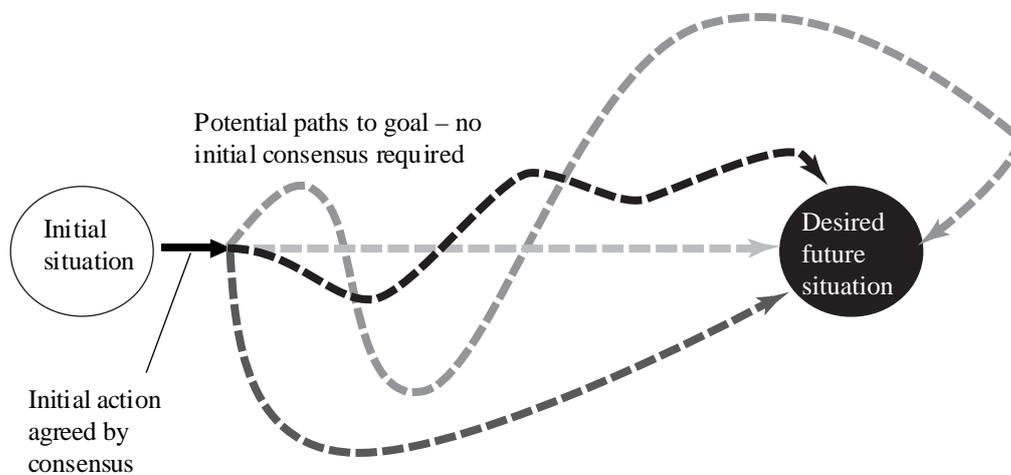
Capacity is an emergent property that evolves partly through the pushes and pulls of contextual factors including global economic trends, national governance, the legacy of regional history and many others. The capacity of an organization derives much of its character from its interaction with the bigger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

Capacity strengthening interventions have a chance of affecting development only to the extent that they are adequately in concert with ongoing development processes. Their effectiveness depends also on whether or not 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 (2005: 20) 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.

Figure 2.1 'Feeling the way to the goal' (Bakewell and Garbutt, 2005: 20)



In *A theory of social change* (2007), Doug Reeler uses multiple diagrams that complement the Bakewell and Garbutt concept shown here (and develops it further by types of change). It is worth noting that in addition to the emergent pathways, the 'desired future situation' is also dynamic or changing in Reeler's models – which we believe more accurately reflects emergent development – whereas it appears fixed in this diagram. An attractive feature of this model, 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 of keeping different actors on track, ensuring the planned activities take place, measuring the outputs, and using 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 their potential for having an impact on development. If CD is a long-term process through uncharted waters, then more navigators and navigating processes are needed (as opposed to simply training up more expert project managers) to adjust course and keep moving forward over time. Although we are keen to receive 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 and 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 programme evidence (for example, surveys and reports) 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 (that is, measurement against a static plan) but 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 potentially forgone.

Attribution

This dilemma is closely related to the previous one in that it also assumes that multiple factors affect the eventual impact, 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 (James, 2001: 12).

Attribution in education programmes has long been a particularly challenging issue. So many variables influence the 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 the difficulty of establishing links between capacity, performance and end impact, whether at the individual or the organizational 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 organizational level. These investments in organizational functions and processes would aim to result in an actual change in programme performance and, ultimately, in the lives of the poorest and most marginalized groups. This presents particular challenges for assessing impact, not least whether it is possible to demonstrate a causal link between a particular organizational 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, James and Wrigley, 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 a well-informed path.

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 programmes 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.

The system-wide change that is being strived for requires efforts by and depends on multiple groups on diverse fronts; hence the merit of attributing impact is highly questionable. The process and multidimensional nature of ... social change means that efforts intertwine in changing contexts,

goalposts inevitably shift, and impact is perhaps best described in terms of 'emergent' phenomena of change. This makes it irrelevant to talk in terms of attribution to specific individuals, efforts or organizations and trying to disentangle which efforts have made what difference. Recognizing the broad system interactions needed for...social change means letting go of an attribution obsession (Guijt, 2007: 20).

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 [downward] accountability" (LaFond and Brown, 2003: 5). This is a debate about M&E and its use in development in general. The argument, already outlined in this paper, is essentially that externally driven M&E primarily based on accounting to donors does not seem designed to yield learning, and may actually produce counterproductive behaviour. Most authors refer to accounting as 'accountability' requirements, which we believe does a disservice to the broad concept of accountability as this includes heavy responsibility to stakeholders as well – that is, downward accountability. We suggest consistently using the term 'donor accounting' or 'upward accountability', which more accurately describes the narrow concept often is referred to as 'accountability'. There is often a focus on predictable, linear causal chains in logical frameworks, which causes organizations to overlook the complexity of the environments in which they operate and not be open to reading the situation and using adaptive management to find the best path forward:

Moreover, many approaches to impact assessment have tended to be mechanistic, linear, standardized 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)

... [D]espite the emergence of some alternatives to mainstream monitoring and evaluation (M&E) approaches, many of the learning and assessment challenges faced by social change-oriented groups are still uncharted in many ways and remain largely unresolved. Many such organizations resort to mainstream M&E approaches that originated under pressure to show measurable and direct changes. These approaches have proven seriously inadequate when applied to efforts aiming to build capacities and social movements, shifting social norms, and strengthening citizenship and democracy. Furthermore, the almost exclusive focus on accountability to donors has often been to the detriment of self-reflection and internal learning that enhances social change processes and to the detriment of accountability to the grassroots (Guijt, 2007: 2).

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

... 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 can be incompatible with systems thinking and learning, only marginally focused on impacts, and more focused on narrow donor accounting of resources and

activities. Watson highlights that 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' (2006: vii).

We are demonstrating here our belief that rigid, inflexible and pre-determined project frameworks may 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 capability of functioning well in society beyond utilizing 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 of and effort needed for implementing effective learning systems (for example, 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 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 that perhaps an 'amicable divorce' should be considered (Mebrahtu *et al.*, 2007: 139). One response would be for learning systems to be set up separately and for complex M&E systems to be stripped down to the bare minimum of donor accounting (Earle, 2002). Yet whether separate or combined, we need to be clearer about what we expect M&E learning approaches to CD to support.

What would we want from learning approaches to the monitoring and evaluation of capacity development that donor accounting approaches are hard placed to deliver?

Some of the literature on the monitoring and evaluation of CD suggests that a broader 'learning approach' might be better for learning and adaptive management, and ultimately for measuring impact, which can be 'told' through a story rather 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 organizations, 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 best illustrated by 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 a local NGO focused on CD service provision. Following the 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 of 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
<p>Performance area: leverage from strategic alliances</p> <ul style="list-style-type: none"> • Monthly committee meetings are being held to further develop our alliance strategy • Increased participation in key conferences and networks 	<ul style="list-style-type: none"> • “Last week at a conference I met with some of the principals of other teacher training colleges in our state and was able to explain fluently our unique approach to developing a 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.”
<p>Performance area: financial sustainability</p> <ul style="list-style-type: none"> • Developed criteria for selecting opportunities • Increased number of responses to requests for proposals (RFPs) (number of proposals submitted) • Carried out a ‘get out the word’ workshop attended by five major funders in the education sector • Overall funding increased 	<ul style="list-style-type: none"> • “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 five 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 direction 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 (that is, useful and relevant) it becomes easier to see what matters and what supports desired change. In the process, their capacities for ‘evaluative thinking’ are developed, which is fundamental for the monitoring and evaluation of CD (Horton *et al.*, 2003: vi.; Watson, 2006: 26). Watson was referring to ‘outcome mapping’ (Earl, Carden and Smutlylo, 2001).

The *most significant change* (MSC) methodology is particularly relevant in this aspect in that it collects significant change stories which infer a relationship between what an organization and the 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 change leads to an understanding of impact and its drivers, and improvement of performance (Davies and Dart, 2005; Wrigley, 2006).

Table 2.1 shows four evaluation types with different focuses and uses for capacity development evaluation, including learning-focused approaches. 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). 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 the linkages between CD interventions and

social change. 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 the rigorous tracking of outputs, which 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 supporting 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.

Table 2.1 also identifies some methods that could possibly 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.

Table 2.1 Selected evaluation types, and different focuses and uses for capacity development

Evaluation type	Purpose and assumptions	Focus	On learning	Possible methods
General impact evaluation	<ul style="list-style-type: none"> Tells us whether an organization is being effective and 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	<ul style="list-style-type: none"> Although learning for CD is secondary, much can still be 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). 	<ul style="list-style-type: none"> Impact evaluations Outcome mapping
Stakeholder-centred impact evaluation	<ul style="list-style-type: none"> Determines the relevance of organizational capacities from the perspective of key stakeholders and the users of organizational services. Parameters for usefulness and relevance of organizational capacities and their application are defined by the stakeholders themselves, and evaluation reveals the meaning that participants derived from the processes. Focus is on capacities related to service delivery (i.e. technical capacities). 	Impact and learning	<ul style="list-style-type: none"> Both impact and learning are the foci, and the evaluation process itself produces lessons on process as well as impact. 	<ul style="list-style-type: none"> Story telling MSC Customer satisfaction surveys 360 degree audits Empowerment evaluations Appreciative inquiry Outcome mapping
Capacity monitoring	<ul style="list-style-type: none"> An ongoing process for analysing the extent to which an organization is developing, retaining and maintaining its capacities for an 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	<ul style="list-style-type: none"> Has a strong internal learning and organizational development focus, and looks for connections to have an impact in a more emergent manner (and sometimes doesn't at all make a formal attempt to link capacity and performance). 	<ul style="list-style-type: none"> Organizational self-assessments and action plans Benchmarking against well-defined metrics Appreciative inquiry
Action learning and capacity studies	<ul style="list-style-type: none"> To 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 the impacts. Looks for performance breakthroughs and their links to differentiated organizational capacities. Looks at the links between CD activities and the capacities developed. 	Learning	<ul style="list-style-type: none"> Looks to answer questions from multiple angles, such as: <ul style="list-style-type: none"> o What are the most critical capacities for effectively contributing to positive change? 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? 	<ul style="list-style-type: none"> Action research Use of evidence in relation to best practice standards Outcome mapping

In the table above, the possible methods referred to in the far right hand column were taken from a presentation on new trends in CD by Beryl Levinger. This table was also developed on the basis of a brainstorm conversation with 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 at other times 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 unpredictability. As such, a helpful way to penetrate the fog is by using methods that naturally thrive in the haze of complexity and don't need to solve capacity as a 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 improve the possibility that CD interventions will support meaningful change.

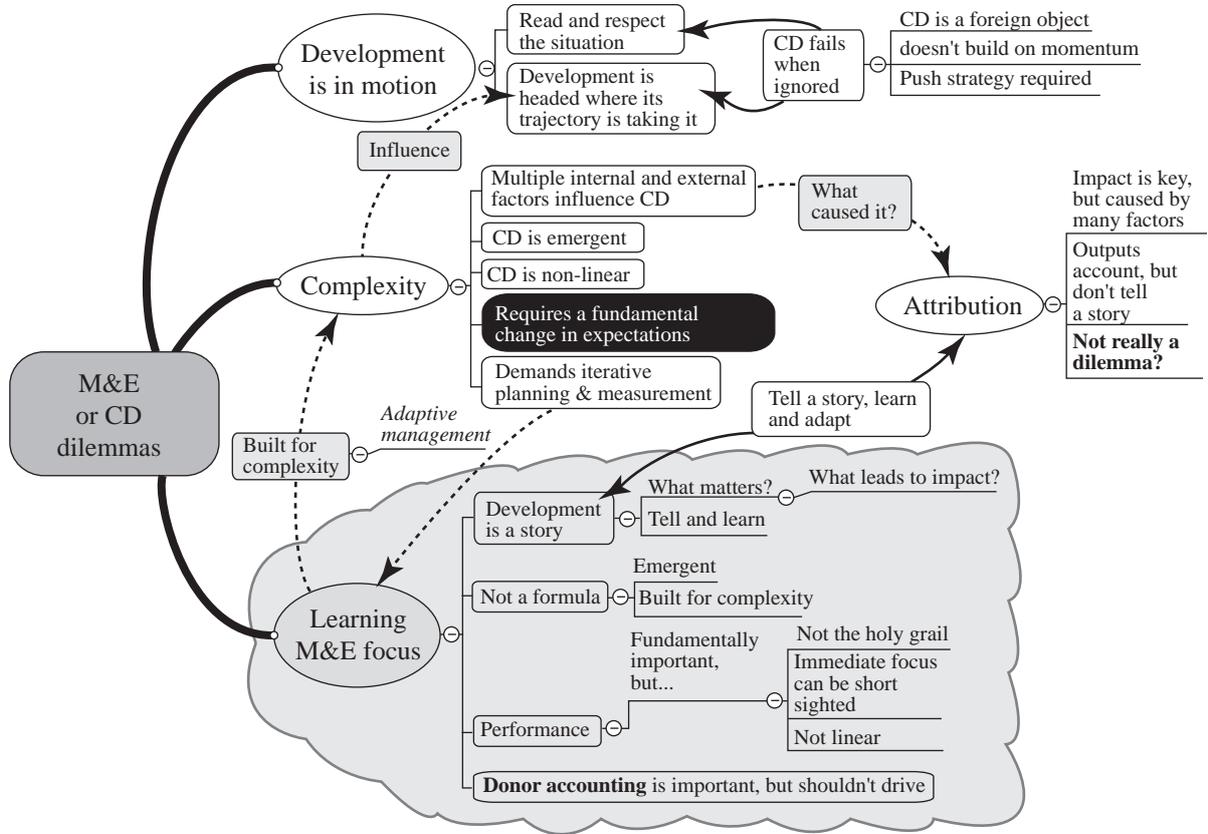
A learning approach to the monitoring and evaluation of CD is intent on helping us get it right, and not only getting it quantified and reported. It is about effectiveness, but also ownership. M&E primarily for donor accounting purposes puts in doubt where the ownership lies, as well as to the very purpose of CD. Contributing to endogenous development in motion assumes that local ownership of development is a desirable condition. The monitoring and evaluation 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 the monitoring and evaluation of CD systems.

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

- Did we thoughtfully analyse the development history and stories before we started, including motivations, traumas, transformational moments, fears, wants and needs?
- 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 – that is, such that there exists an untapped, willing momentum for change?
- Did we frequently debate progress, delays, insights, assumptions and activities?
- 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 and do everything we could to promote positive change?
- 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?

Figure 2.2 summarizes some of the principal dilemmas of the monitoring and evaluation of CD as described in this section:

Figure 2.2 Dilemmas with the monitoring and evaluation of capacity development



3 Concluding thoughts

The discussion around the monitoring and evaluation 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, it is not possible to use monitoring and evaluation to determine if CD efforts are leading us to a better understanding of development and development programming. 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 positive change), and developing or strengthening capacities to result in greater system and organizational readiness and ability. Organizations are not learning systemically and systematically 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, "[i]t 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, 2008: xiii). How many other capacity development interventions continue to be carried out without serious reflection on what works and what does not?

There is an urgent need to improve development programming conception, design, implementation and learning in general. The unique nature of capacity development makes it the ideal 'learning laboratory' for discovering key insights into development and the links between capacity, performance and impact. The CD that results from practising '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 programme implementation. Such effects can be quite pronounced, as when the process of clarifying the programme'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 "[t]he 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 organizational culture oriented to performance, learning and self-reflection" (Morgan, 1999: 14). Organizations 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, and complexity and emergence as opposed to linear development thinking.

... no one is in overall control of what is happening, and although patterns of relating tend in a particular direction, the exact global pattern that emerges is unpredictable (Mowles, Stacey and Griffin, 2008: 810).

Because of the scale and complexity of the game being played by these [social] actors, it can only result in unpredictable and unexplained consequences no matter how clear and logical the strategy pursued by any actor (Mowles, *et al.*, 2008: 815).

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 the introduction 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 better ways to complement and strengthen development processes and individual, organizational and system response-ability over time. Although the task is daunting, the current and largely ineffective status quo needs to be questioned actively and alternative approaches tested. We believe that the following recommendations are important for improved monitoring and evaluation of capacity development:

- a) incorporation of organizational learning approaches to the monitoring and evaluation of CD;
- b) large-scale experimentation and action research;
- c) use of theory of change approaches for designing systems for monitoring and evaluating CD.

Incorporation of organizational learning approaches to the monitoring and evaluation of capacity development

We have already established throughout the book the importance of learning approaches to monitoring and evaluation for capacity development at the level of the project, activity or intervention, and as such have not repeated that in this sub-section. Although it is clear that the required shift for CD to organizational learning is more a cultural than procedural one, something needs to be done to start the process, and to test whether learning-based organizational cultures translate into a more effective conception, design, implementation and evaluation of capacity development programming.

The South African NGO Community Development Resource Association's (CDRA) 'homeweek' is an interesting organizational learning model. It is a dedicated monthly "week-long process of organizational connecting, strategizing, 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 organizational 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 organizing 'structure-in-motion'. Recognition was given that all organizational 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 organizational and maintenance activities take place, including learning (Soal, 2001).

These activities are wide-ranging, including the sharing of experiences by different members of the organization, some of them in very creative ways; strategy and planning meetings; business meetings; and the 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 organizational” (Soal, 2007: 6).

Soal explains the transformative nature of this type of practice: “We have seen that organizations 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” (Soal, 2007: 2).

Another example of the 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, Mancini and Guijt, 2006: 137). “ALPS was not designed to feed the insatiable desire for upward reporting” (David, *et al.*, 2006: 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 (David, *et al.*, 2006: 140).

ALPS is based on three core elements: principles, attitudes and behaviours, and organizational policies and processes. All of 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 the excluded people with whom Action Aid and its partners work. These elements connect with all of Action Aid’s organizational processes, including appraisals; strategies; strategic plans; annual plans and budgets; strategic reviews; peer reviews; organizational climate 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” (David, *et al.*, 2006: 145). These changes include improved decision making by virtue of getting the 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 (David, *et al.*, 2006: 143), they offer key insights into 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 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 are, however, 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.

We recognize 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 organizations 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 gradually to introduce this also within public sector organizations, where they are forming partnerships and collaborations.

Large-scale experimentation and action research

The experimentation recommended at the organizational level also needs to be taken to a broader system level if the monitoring and evaluation of CD is to have the wider impact that we have argued it should.

At best, we imagine a rather gradual progress towards an eventual tipping point when the monitoring and evaluation 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 the monitoring and evaluation of 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 're-imagine' 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 mobilize 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 practised, by: promoting empowering relationships; supporting rallying ideas; mobilising dynamic agents; 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 is 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 examples of donors supporting informal monitoring using a systems thinking approach" (referring specifically to the ECDPM case studies) (Watson, 2006: viii), Watson points out:

There is persuasive evidence of the value and effectiveness – in contributing to organizational 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 (Watson, 2006: viii).

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 the monitoring and evaluation 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 the monitoring and evaluation of CD and can help answer critical questions such as:

- What are the most critical capacities for organizational effectiveness?
- What are the characteristics of organizations that are most likely to generate innovation and catalyse broad-scale change?
- Do high capacity organizations working together create a 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 programmes the most effective way for increasing needed capacities at different levels, or would more targeted social innovation investment yield better results?

Use of theory of change approaches for designing the monitoring and evaluation of capacity development systems

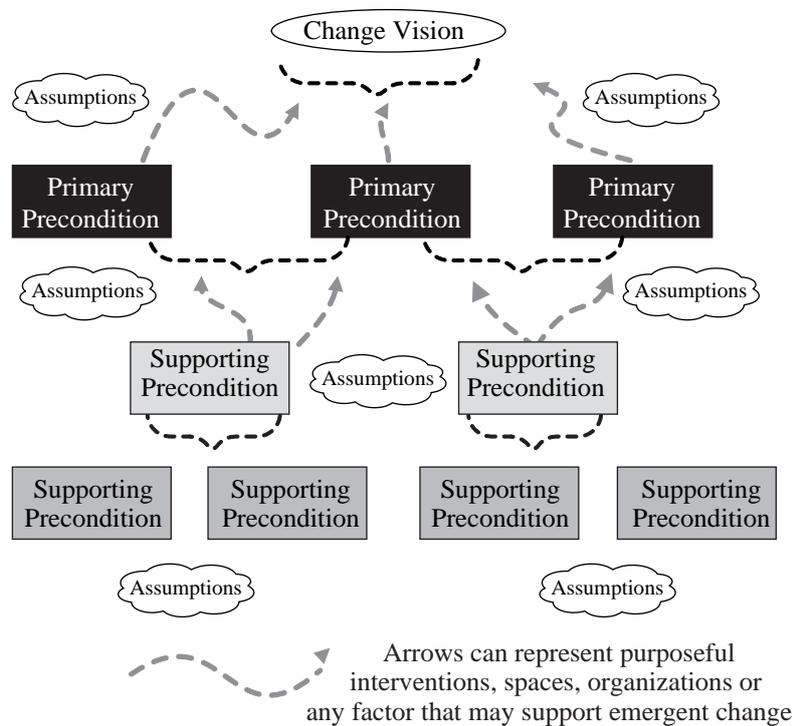
What can a theory of change offer?

[A]ny process for intervening sensibly in real-world situations to bring about ‘improvement’ must have some ideas – some theory – about the nature of social reality, whether it is made explicit or not. There must be some theory which makes any chosen process of intervention ‘sensible’ (Checkland and Poulter, 2006: 172).

Development practice is informed by theories of change, but individuals and organizations may not make them explicit. Practitioners may be unaware of the extent to which strategic choices and debates are informed by disparate thinking about how history happens and the role of purposeful intervention for progressive social change (Eyben *et al.*, 2008: 201).

Theory of change (TOC) represents an organization’s understanding of how development (change) happens with regard to the issues that it exists to address (mission or purpose). This includes a particular understanding of the demands or needs of the primary stakeholders (that is, purpose) and the conditions that are needed to support the emergence of change, given the complexities (for example, power, culture, systems and other actors) in the broader environment. TOC principles can be used as part of a planning process for visualizing, constructing and substantiating the elements, conditions and interventions that are fundamental for supporting positive change. A key advantage of TOC is that it encourages the design of a programmatic intervention model that includes a broad, systemic look at the conditions that can support the emergence of change in a programme, project or overall organizational strategy (not only the areas that a particular project might typically take into account). Most importantly, it encourages the identification of underlying assumptions that, when visible, can strengthen organizational debate and analysis but when hidden, can lead us to ‘blind prescriptions to change’.

Figure 3.1 Simplified TOC diagram (adapted from ActKnowledge, 2009)



There is a need to observe and understand the change processes that already exist in a living social system. If we can do this before we rush into doing our needs analyses and crafting projects to meet these needs, we may choose how to respond more respectfully to the realities of existing change processes rather than impose external or blind prescriptions based on assumed conditions for change (Reeler, 2007: 2).

Visualizing these conditions and the inherent complexities within, it becomes more feasible to analyse whether the approaches and interventions that a particular organization chooses to implement, in addition to its assumptions of the roles and interventions of other actors, are systemically well thought out in support of the broader positive change vision. This can also reveal gaps in a particular CD project design – both at the level of conditions and interventions necessary for supporting emergent change thoughtfully – and makes it clearer to see what needs to be measured. This may sound a recipe for visual messiness, but as the example in *Figure 3.1* illustrates, the TOC can be helpful in visualizing how we think change happens and what we should do about it.

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 (Keystone, n.d.).

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 (Keystone, n.d.).

Box 1. A clarifying note on TOC diagram structure and non-linearity

Conditions and interventions at lower levels of a TOC do not *cause* higher level conditions to occur – that is, there is not a linear, cause-effect relationship. One thing can be said to cause another

... if the cause is both necessary and sufficient for its effect. One thing is necessary for another if the other cannot occur unless the first one does. One thing is sufficient for another if the occurrence of the first assures the occurrence of the second (Ackoff, 1999: 10).

Lower level preconditions are necessary ‘conditions’ that support higher level preconditions, but they are never sufficient for their occurrence because all conditions are emergent – that is, they have properties which are more than the sum of their parts (Flood, 2001: 133) and which are the result of multiple factors that complexity renders ‘inherently unknowable to the human mind’ (Flood, 1999: 86). As such, interventions are ultimately only part of a myriad of factors that might contribute to overall change.

Emergence is an unplanned and uncontrollable process in which properties such as capacity emerge from the complex interactions among all actors in the system and produce characteristics not found in any of the elements of the system (Land, Hauck and Baser, 2009: 2).

The path between preconditions is, like development in general, non-linear and the TOC can be presented in multiple ways, including more creative organic looking diagrams that are clearly non-linear. If, however, an organization’s overall perception of change is predominantly linear, then it is indeed possible that this orientation may manifest itself through the way the TOC is articulated.

As an example, the NGO Pact developed and tested a TOC framework for the Zambian HIV/AIDS Learning Initiative (ZHLI), a programme focused on strengthening the operational, technical and financial capacity of Zambian NGOs, networks and ISOs leading multisectoral HIV and AIDS prevention and care activities (this report can be found at www.impactalliance.org). In its TOC-based M&E framework, the project intervention model was mapped. This showed how programme objectives and components, working with targeted actors and institutions, should plausibly effect change types (that is, collaboration and alliances, culture, learning and innovation, policy, process and systems, strategy, structure and technology) – which should lead to broader HIV and AIDS service reach and quality. This included a look at how other actors relate to the programme interventions and a story-based approach for collecting and validating programme 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 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.

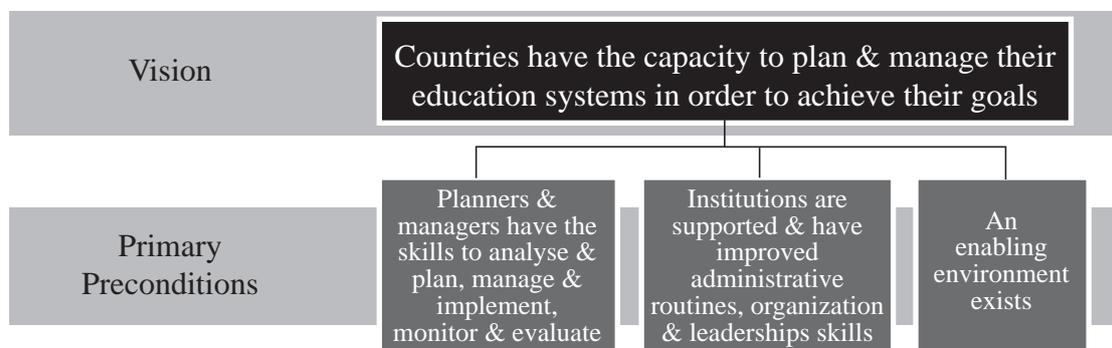
Hypothetical example of theory of change use in Education for All

IIEPs mission is “to strengthen the capacity of countries to plan and manage their education systems” (IIEP, 2009) and it does this by:

Training planners and managers in skills to analyse 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 cooperation and networking (IIEP, 2009).

On a basic, broad level the following diagram shows IIEP's vision for change and preconditions for change, without interventions included:

Figure 3.2 IIEP change vision



This basic model is essentially a partial theory of change (TOC) in that it includes a vision (though the existence of higher levels of vision – something that translates into quality of life for individuals, communities and so on – isn't necessary for the purposes of this example) it identifies the preconditions for success that must be in place before the vision can become possible, and by adding interventions – which IIEP defines as capacity development in all of its core documents – IIEP will posit how it sees development happening – that is, how it sees these conditions coming about, at least from the CD lens.

This is a shared TOC (shared by 'default') in that IIEP is not the only organization that assumes the challenge of working towards country capacity to plan and manage education systems. Therefore, from the outset these preconditions include the potential results of IIEP *and* the results of other organizations and actors (the governments themselves as well as other development and CD service providers) that are among the many factors that may contribute to the emergence of the vision in this TOC.

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.

Utilizing this broad TOC framework, we offer some observations and questions regarding the links between the monitoring and evaluation 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	<ul style="list-style-type: none"> • 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)?
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 that IIEP chooses to pursue are the most effective ways to catalyze this TOC. The answer to these questions should take into account IIEP's 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	<ul style="list-style-type: none"> • Do IIEP's CD programme and project designs utilize the most effective interventions for achieving the desired impact? • Does IIEP help build resilience, readiness and ability in the governments, organizations and individuals that it strengthens? • To what extent do IIEP's CD efforts take into account complex adaptive systems, including influences and relationships between levels of capacity (e.g. individual, organization and system)? • Does IIEP adequately utilize learning-based approaches to M&E?
Description	
This is in reference to the level of programme and project interventions – that is, the operationalizing of IIEP's programmatic strategies. At this level, the monitoring and evaluation of CD should help answer whether IIEP selects the most adequate CD interventions to get to impact. In other words, do training, strengthening administrative routines, and so on, effectively achieve the desired impacts? Additionally, the monitoring and evaluation should inform IIEP whether it builds adequate response-ability (that is, standing capacities), including adaptive management abilities that help member states, and support organizations and individuals in managing education priorities and relationships in complex, constantly adapting systems.	

Partner level

M&E of CD objective	Core questions
Learn from partner experiences and continually define the most ideal partners for responding to EFA	<ul style="list-style-type: none"> • 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 the areas of our TOC that are currently underserved?
Description	
As part of a strategic partnership management, the monitoring and evaluation of CD should aid IIEP in gathering lessons learned from implementing partners (governments, NGOs and others) to help inform its decision-making at different levels. Additionally, analysing M&E results at a cumulative level can help determine if IIEP 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 a good option for understanding how intangibles such as relational and adaptive management capacities need to be developed from an open system perspective informed by learning-based monitoring and evaluation for CD.

Conclusion

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

Bibliography

- Ackoff, R. 1999. *Ackoff's best: his classic writings on management*. John Wiley & Sons Inc.
- ActKonwledge. 2009. Retrieved June 2008 from:
<http://theoryofchange.org/>
- 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*. Washington, DC: Pact.
- Bloom, P.N.; Dees, G. 2007. "Cultivate your ecosystem". In: *Stanford Social Innovation Review*, 47-53.
- Buckingham, M.; Coffman, C. 1999. *First, break all the rules: what the world's greatest managers do differently*. New York: 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.
- Checkland, P.; Poulter, J. 2006. *Learning for action: a short definitive account of soft systems methodology and its use for practitioners, teachers and students*. Chichester: John Wiley & Sons, Ltd.
- Collins, J. 2005. *Good to great and the social sectors*. Boulder, CO: Jim Collins.
- Community Development Resource Association (CDRA). 2001. *Making the learning organisation literal CDRA's homework*. Retrieved 16 September 2008 from:
www.cdra.org.za/nuggets/CDRAs%20homework%20%20by%20Sue%20Soal%20%20October%202001%20CDRA%20Nugget.htm
- Crutchfield, L.R.; Grant, H.M. 2008. *Forces for good: the six practices of high-impact nonprofits*. San Francisco, CA: 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*. Retrieved June 2008 from:
www.mande.co.uk/docs/MSCGuide.pdf
- Dlamini, N. 2007. Stories and M&E. In: ONTRAC. Oxford: 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*. Oxford: INTRAC.

- Eyben, R.; Kidder, T.; Rowlands, J.; Bronstein, A. 2008. "Thinking about change for development practice: a case study for Oxfam GB". In: *Development in Practice*, 18, 12.
- Flood, R.L. 2001. "The relationship of 'systems thinking' to action research". In: P. Reason, H. Bradbury (Eds.), *Handbook of action research: participative inquiry and practice*. Thousand Oaks, California: Sage.
- Fowler, A. 2000. *The virtuous spiral: a guide to sustainability for NGOs in international development*. London: Earthscan.
- Guijt, I. 2007. "Assessing and learning for social change". In: IDS (Ed.), *IDS Discussion Papers*. Brighton: IDS.
- Hailey, J.; James, R.; Wrigley, R. 2005. "Rising to the challenges: assessing the impacts of organizational capacity building". In: INTRAC (Ed.), *Praxis Papers*. Oxford: INTRAC.
- 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.; Perez, J.; Chang, M.S.; Vernoooy, R.; Watts, J. 2003. *Evaluating capacity development: experiences from research and development organizations around the world*. The Netherlands: International Service for National Agricultural Research (ISNAR); Canada: International Development Research Centre (IDRC); the Netherlands: ACP-EU Technical Centre for Agricultural and Rural Cooperation (CTA).
- Independent Evaluation Group (IEG). 2008. *Using training to build capacity for development: an evaluation of the World Bank's project-based and WBI training*. Washington, DC: The International Bank for Reconstruction and Development/World Bank.
- International Institute for Educational Planning (IIEP). 2009. *Our mission*. Retrieved May 2008 from:
www.iiep.unesco.org/aboutiiep/about-iiep.html
- 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*. Cape Town: CDRA.
- Keystone. No date. Retrieved May 2008 from:
www.keystoneaccountability.org/node/34
- 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.
- Land, T.; Hauck, V.; Baser, H. 2009. "Capacity development: between planned interventions and emergent processes. Implications for development cooperation". In: ECDPM (Ed.), *Capacity Change and Performance, Policy Management Brief*. Maastricht: ECDPM.
- Lessik, A.; Michener, V. 2000. "Measuring institutional capacity". In: USAID CDIE, (Ed.) *Recent Practices in Monitoring and Evaluation TIPS*. Washington, DC: USAID.
- Linnell, D. 2003. *Evaluation of capacity building: lessons from the field*. Washington, DC: Alliance for Nonprofit Management.
- Lusthaus, C.; Adrien, M.H.; Perstinger, M. 1999. *Capacity development: definitions, issues and implications for planning, monitoring and evaluation*. Montreal: Universalia.

- Lusthaus, C.; Anderson, G.; Murphy, E. 1995. *Institutional assessment: a framework for strengthening organizational capacity for IDRC's research partners*. Ottawa: International Development Research Centre.
- Lusthaus, C., Adrien, M.H., Anderson, G., Carden, F. and Montalvan, G.P. 2002. *Organizational assessment. A framework for improving performance*. Ottawa/Washington DC: International Development Research Centre and Inter-American Development Bank.
- Mebrahtu, E.; Pratt, B.; Lönnqvist, L. 2007. *Rethinking monitoring and evaluation*. Oxford: INTRAC.
- Morgan, P. 1997. *The design and use of capacity development indicators*. Gatineau: CIDA.
- Morgan, P. 2006. *The concept of capacity*. Maastricht: ECDPM.
- Mowles, C.; Stacey, R.; Griffin, D. 2008. "What contribution can insights from the complexity sciences make to the theory and practice of development management?" In: *Journal of International Development*, 20, 804-820.
- Oakley, P.; James, R. 1999. *Monitoring and evaluation of capacity building: a desk review of UNICEF draft*. Oxford: INTRAC.
- OECD. 2006. *The challenge of capacity development working towards good practice*. Paris: OECD.
- Pasteur, K. 2006. Learning for development. In: R. Eyben (Ed.), *Relationships for aid*. London: Earthscan.
- Reeler, D. 2007. *A theory of social change and implications for practice, planning, monitoring and evaluation*. Cape Town: CDRA.
- Roche, C. 1999. *Impact assessment for development agencies: learning to value change*. Oxford: Oxfam.
- Sen, A. 1999. *Development as freedom*. New York: Anchor Books.
- Soal, S. 2001. Making the learning organisation literal: CDRA's homework. *CDRA Nugget* Retrieved 16 September 2008 from:
www.cdra.org.za/nuggets/CDRAs%20homework%20%20by%20Sue%20Soal%20%20October%202001%20CDRA%20Nugget.htm
- 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". In: *GUNI – Higher Education for Human and Social Development: new roles and emerging challenges (Palgrave)*, 12.
- Taylor, P.; Clarke, P. 2008. *Capacity for a change*. Brighton: IDS.
- UNDP. No date. *The Human Development concept*. Retrieved 1 June 2009 from:
<http://hdr.undp.org/en/humandev/>
- Vansant, J. 2000. *A composite framework for assessing the capacity of development organizations*. Prepared for USAID by J. Vansant. Retrieved 1 June 2009 from:
www.ngomanager.org/vansantarticle.htm

Bibliography

- 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*. Oxford: INTRAC.

IIEP publications and documents

More than 1,200 titles on all aspects of educational planning have been published by the International Institute for Educational Planning. A comprehensive catalogue is available in the following subject categories:

Educational planning and global issues

General studies – global/developmental issues

Administration and management of education

Decentralization – participation – distance education – school mapping – teachers

Economics of education

Costs and financing – employment – international cooperation

Quality of education

Evaluation – innovation – supervision

Different levels of formal education

Primary to higher education

Alternative strategies for education

Lifelong education – non-formal education – disadvantaged groups – gender education

Copies of the Catalogue may be obtained on request from:

IIEP, Publications and Communications Unit

info@iiep.unesco.org

Titles of new publications and abstracts may be consulted online:

www.iiep.unesco.org

The International Institute for Educational Planning

The International Institute for Educational Planning (IIEP) is an international centre for advanced training and research in the field of educational planning. It was established by UNESCO in 1963 and is financed by UNESCO and by voluntary contributions from Member States. In recent years the following Member States have provided voluntary contributions to the Institute: Australia, Denmark, India, Ireland, Netherlands, Norway, Spain, Sweden and Switzerland.

The Institute's aim is to contribute to the development of education throughout the world, by expanding both knowledge and the supply of competent professionals in the field of educational planning. In this endeavour the Institute co-operates with training and research organizations in Member States. The IIEP Governing Board, which approves the Institute's programme and budget, consists of a maximum of eight elected members and four members designated by the United Nations Organization and certain of its specialized agencies and institutes.

Chairperson:

Raymond E. Wanner (USA)

Senior Adviser on UNESCO issues, United Nations Foundation, Washington DC, USA.

Designated Members:

Christine Evans-Klock

Director, ILO Skills and Employability Department, Geneva, Switzerland.

Carlos Lopes

Assistant Secretary-General and Executive Director,

United Nations Institute for Training and Research (UNITAR), United Nations, New York, USA.

Jamil Salmi

Education Sector Manager, the World Bank Institute, Washington DC, USA.

Diéry Seck

Director, African Institute for Economic Development and Planning, Dakar, Senegal.

Elected Members:

Aziza Bennani (Morocco)

Ambassador and Permanent Delegate of Morocco to UNESCO.

Nina Yefimovna Borevskaya (Russia)

Chief Researcher and Project Head, Institute of Far Eastern Studies, Moscow.

Birger Fredriksen (Norway)

Consultant on Education Development for the World Bank.

Ricardo Henriques (Brazil)

Special Adviser of the President, National Economic and Social Development Bank.

Takyiwaa Manuh (Ghana)

Director, Institute of African Studies, University of Ghana.

Philippe Méhaut (France)

LEST-CNRS, Aix-en-Provence, France.

Xinsheng Zhang (China)

Vice-Minister of Education, China.

Inquiries about the Institute should be addressed to:

The Office of the Director, International Institute for Educational Planning,
7-9 rue Eugène Delacroix, 75116 Paris, France

The book



Monitoring and evaluation of capacity development faces a fundamental dilemma. On the one hand, capacity development is a long-term process which occurs through a series of changes, related to values, attitudes and norms, which are difficult, if not impossible, to measure. On the other hand, evaluation tends to focus on short-term changes and to examine what is easy to perceive. As a result, at times we privilege short-term efforts which lead to visible change, even if they may be detrimental to sustainable improvement. However, when evaluation tends to focus on long-term impact, measurement and attribution difficulties quickly become evident.

Although there is no easy solution to this, some new approaches may succeed better in assessing the impact of capacity development. They are characterized by stronger participation by beneficiaries in the assessment of change. They use a mixture of qualitative and quantitative methods. They focus on the relevance of interventions in specific contexts. Most importantly, they want the monitoring and evaluation process to be a form of capacity development; in other words, participants learn through and from the evaluation about how the programmes and their own practices can change for the better.

The authors

Peter Taylor researches, teaches and advises on the theory and practice of capacity development, engagement of education systems and institutions in participatory and social change, and facilitation of learning in development contexts. He has recently joined the IDRC, Canada, as a Senior Programme Specialist after 6 years at the Institute of Development Studies, UK.

Alfredo Ortiz focuses, through teaching, consulting and action research, on systemically understanding and strengthening the links between organizational capacity, programmatic strategy and development impact – primarily of NGOs. He teaches at the Monterey Institute of International Studies, has worked for ten years as a consultant in NGO capacity development, and currently studies at the Institute of Development Studies, UK.



United Nations
Educational, Scientific and
Cultural Organization



International Institute
for Educational Planning