www.elsevier.com/locate/worlddev

CrossMark

Capacity Development Evaluation: The Challenge of the Results Agenda and Measuring Return on Investment in the Global South

BERTHA VALLEJO [a,b] and UTA WEHN [c,*]

[a] Tilburg University-CentER, The Netherlands
[b] UNU-MERIT, Maastricht, The Netherlands
[c] UNESCO-IHE, Delft, The Netherlands

Summary. — This study reviews the evaluation of capacity development, identifying capacity development (CD) modalities and the schools of evaluation currently in place. The research joins the results agenda debate, arguing that in dealing with CD interventions, pre-defined indicators fail to represent the process and the key elements that take CD recipients toward patterns of change. The study highlights the fact that CD deals with projects that, by their nature (consisting of change processes designed to initiate change in people, organizations, and/or their enabling environment), rely more on non-planned changes than on the pre-defined indicators and results to contribute to livelihood improvements and social transformation. The study recognizes the difficulty of evaluating CD under straight-forward mechanisms. It concludes that the existing approaches are not adequate to truly capture or measure impact, as CD projects, restricted by previously agreed budgets, resources, and time frames, are usually not designed to evaluate the sustainability of change and its impact over the medium or long term. As resources are scarce, donor agencies and policy-makers need to know the value of CD in order to best prioritize their investments. However, due to the nature of these projects, measuring the return rate between the project cost and its impact remains a difficult task. There is a need for new, multi-path approaches to capturing changes in capacity in order to serve as a basis for decision-making regarding CD investments.
© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. INTRODUCTION

The role and nature of development cooperation was the topic of various High Level Fora (HLF) on Aid Effectiveness in Rome (2003), Paris (2005), Accra (2008), and Busan (2011) that resulted in global commitments and development practice principles. The latest High Level Forum in Busan in 2011 led to *The Global Partnership for Effective Development Co-operation* (OECD, 2011). Throughout this ongoing debate, capacity development (CD) has been the key priority of development cooperation (Pearson, 2011). Moreover, CD is now an explicit and integral part of the Sustainable Development Goals (SDGs), with their focus on implementation aspects; compared to the Millennium Development Goals (MDGs) (Joshi, Hughes, & Sisk, 2015), this represents a paradigmatic shift in the emphasis placed on capacity development, now considered as a vehicle for meeting the SDGs and for sustaining these achievements in the long term (Wehn, 2014a).

CD has high financial relevance, not only for the recipient country—where in many cases it is the most important component of the developmental efforts—but also for the donors and implementing agencies, as it is a core function of international development organizations (e.g., UN, OECD, DFID, USAID). In 1996, the United Nations, through its UN General Assembly Resolution A/RES/50/120 Article 22, declared capacity development as an essential path through which development occurs (and as an essential part of the operational activities of the United Nations system at the country level), and not only as a strategy for development.

The need for CD financial inflows varies among the recipient countries. While middle-income or emerging economies

are able to self-finance most of it, the poorest or low-income countries rely in about 25% of their development investments on these money flows, in the form of grants or highly concessional loans (Guicquero, 2015). Although development assistance financing increased considerably during the last decade (European Commission, JRC-IPTS, & Joanneum Research, 2002; Morgan, 1999; OECD, 2007–2012; Raynard, 2000; Smillie, 1995) and will continue increasing in overall terms, the share of aid going to the poorest countries in the post-2015 agenda is not yet defined (Gunzburg, 2015).

As money has become a scarcer input and taxpayers are demanding clearer value for money explanations, it is not surprising that debate among development actors is increasingly focusing on project results and rates of return on the amounts invested. A clear example of this value for money or results focus is the recent Independent Commission for Aid Impact (ICAI)'s report on UK Official Development Assistance (ODA)'s expenses (Hencke, 2015; Valters, 2015). Simultaneously, there is also an increasing debate over the performance of development agencies and NGOs in achieving long-term socio-economic transformation goals (Banks, Hulme, & Edwards, 2015; Green, 2015; Jozwiak, 2015). In this context, the quantifiable outcomes of infrastructure investments and predetermined quantified results are typically more easily measured and more highly valued by development agencies than the more intangible or "soft" outputs of capacity development interventions (Roberts, 2013). Yet the main difficulty in quantitatively measuring the return rate for the money invested in CD

---

interventions is precisely the nature of these types of projects, as they are developmental and concerned with organizational, social, and individual changes (Blume, Ford, Baldwin, & Huang, 2010; Preskill & Boyle, 2008) aimed at livelihood improvements and social transformation. Due to the nature of these changes, it is difficult to appreciate in the short run what knowledge and skills are adopted by participants of CD interventions once the external partner support and money have left.

Studies exploring the sustainability patterns after the completion and withdrawal of CD interventions, resources, and expertise are almost nonexistent (Themistocleous & Wearne, 2000). This is not surprising, as in practice, the budget allocated to a specific project does not include monitoring once the project has been completed. Experience shows that once the donor and implementers complete the project and leave the country or organization in which the project was implemented, the achieved results are rarely sustained in the medium and long run (Clark, Hall, Sulaiman, & Naik, 2003; Godfrey et al., 2002; Pascual Sanz, Veenstra, Wehn de Montalvo, van Tulder, & Alaerts, 2013). In practice (in many cases), not only do the external partners move out, but also the local counterparts move on to other local projects (as participation in CD projects is an important budget component for both local participants and the local organizations). Therefore, without a budget allocated to follow-up and without proper ownership and integration of the learning goals by the local counterpart, there are no resources to be allocated to continuity.

The evaluation of CD interventions and their modalities is a very complex task, as the intangible effects (i.e., social and individual transformations) are not easily grasped by commonly adopted evaluation methods (Blume et al., 2010; Preskill & Boyle, 2008). In seeking to fulfill the requirements of the donors, not only has the focus of CD projects or implementations evolved over time, but so have the modalities adopted by the implementers in each intervention, as well as the monitoring and evaluation techniques used in their assessment (UNDP, 2013). However, "knowing" and "doing" diverge many times, as changes in habits and practices take time (Argyris & Schön, 1978; Kolb, 1984; Senge, 1990).

This paper aims to contribute to the ongoing "results agenda" and "value for money" debate by emphasizing the importance of understanding the complexity of CD, the relevance of non-planned results, and the fact that CD is not a simple transfer mechanism of know-how but rather a process, which in partnership with the recipients develops the learning behind social transformation through the improvement of livelihoods and the capacity to adapt this learning (or know-how) to different, and specific, contexts, cultures, and realities. It is a process that enables the recipient to do things differently and to modify habits and practices. Therefore, the focus is on the doing, rather than on the knowing how.

The research presents a review of the literature on the main elements behind the most common evaluation methods of CD interventions. It identifies and highlights commonly adopted capacity development interventions and modalities as well as the schools of evaluation currently in place, along with their main elements and characteristics. The study concludes with an assessment of evaluation practices commonly adopted in CD projects and the recognition of the difficulty of evaluating CD interventions under straightforward mechanisms due to the nature of both capacity and capacity development projects.

The paper is organized as follows: Section 2 presents a description of capacity development terminology and its evolution over time. Section 3 presents a description of capacity development interventions and the most commonly adopted modalities. Section 4 discusses the evaluation of capacity development. Section 5 presents evaluation methods used to evaluate capacity development efforts. Section 6 concludes the paper with a discussion of the findings.

2. CAPACITY DEVELOPMENT IN PERSPECTIVE

The common goal of development projects through the years has been poverty alleviation and livelihood improvement for the local people (WRI, 2008). The paradigms followed by development agencies in achieving these objectives have evolved over time, shifting from supply-driven technical assistance based on inputs toward demand-driven capacity development based on outcomes, and they have been strongly influenced by the Millennium Development Goals (UNDP, 2009).

In the 1950s and 1960s, developmental aid targeted institutional capacity building through technical assistance, which aimed at providing local public institutions with the finances and physical infrastructure required to manage programs of public investment (Lusthaus, Adrien, & Perstinger, 1999; Sastre Merino & de los Rios Carmenado, 2012). In the 1960s and 1970s, the focus shifted to institutional development and strengthening (Lusthaus et al., 1999; Sastre Merino & de los Rios Carmenado, 2012). In the 1970s and 1980s, the focus of development shifted toward the people, stressing the importance of intangible aspects, such as education, health, and population through the development of human resources including knowledge, skills, and attitudes (Enemark & Ahene, 2002). The emerging discourse on knowledge societies facilitated by the rapid and wide diffusion of information and communication technologies (Mansell & Wehn, 1998) brought the importance of knowledge for development to the fore.

Institutional economists resurged in the 1980s and 1990s, with an emphasis on the major stakeholders: government, non-governmental organizations (NGOs), and private organizations, as well as their networks and external environment (Lusthaus et al., 1999). The focus of this approach is on economic behavior and sustainability.

In the late 1980s, the concept of capacity development emerged in the literature, evolving from years of development interventions between North and South countries (Lusthaus et al., 1999) and embedded in the systems perspective framed by the fundamentals of evolutionary economics (Edquist, 1997; Nelson & Winter, 1982).

Table 1 presents a review of the four main interrelated approaches to CD identified by the literature: organizational, institutional, systems, and participatory.

This study understands capacity development as the process through which individuals, groups, organizations, institutions, and societies increase their abilities to: (i) perform core functions, solve problems, define and achieve objectives; and (ii) understand and deal with their development needs in a broad context and in a sustainable manner (UNDP, 1997; UNESCO, 2006, Chapter 3). Under this approach, CD is an umbrella concept connecting elements from other developmental approaches with a long-term, demand-driven perspective, seeking social change through sustainable social and economic development (Alley & Negretto, 1999; Lusthaus et al., 1999; Morgan, 1998).

The CD approach, as understood in this research, suggests enhancing and strengthening existing capacities, not building them from scratch (Dia, 1996; Ohlbaum, 2015a) as had arguably been implied by the term "capacity building" used in previous decades. It understands development as an endogenous transformation process undertaken by LDCs and developing countries and supported, not steered, by external interventions (Kaplan, 2000; Kuhl, 2009).

Table 1. *Four approaches to capacity development*

Approach	Focus	Strengths	Difficulties
Organizational	Focuses on identifying the components of capacity within an organization (Morgan, 1989). It sees the system through the eyes of an organization	It has much in common with the field of organizational theory and change. The unit of change is clear	It is not clear how to change organizations in the developing world. Organizational change is necessary but not sufficient (on its own)
Institutional	It builds the capacity to create, change, enforce, and learn from the processes and rules that govern society (Cohen, 1994). Strengthens and restructures existing local institutions. The emphasis is on individual institutions	At the macro level, allows dealing with norms, cultural values, incentive systems, and beliefs, which underlie development problems	The boundary between institutional change and CD is unclear
Systems	It is a dynamic, multidimensional, holistic, and interrelated approach that covers the institutional and organizational perspectives. It suggests that CD should build on existing capacities, rather than build new systems (OECD, 1996; UNDP, 1997).	It is comprehensive, flexible, and emphasizes linkages between elements	The biggest difficulty is to identify what is and what is not a CD activity. It is unclear where a system change effort starts
Participatory	A process approach that emphasizes the relevance of the means used to achieve CD. It is people-centered [in the sense of organization] and non-hierarchical. It empowers partnership and ownership as key elements for CD (Fowler, 1997). This approach overlaps with aspects from the other three approaches presented here (Morgan, 1998)	It has a clearly defined scope and shares concepts of development, emphasizing participation, ownership, and power sharing	Interventions with a narrow development outcome are easily confused as CD. However, interventions that do not contribute to building capacity should not be considered CD

Source: Adapted from Lusthaus *et al.* (1999) and Kuhl (2009).

This type of approach is increasingly gaining support by practitioners in CD, as Diana Ohlbaum, former deputy director of USAID's Office of Transition Initiatives, states in the Devex (a development media platform) series: "[We] should be supporting change, not driving it. Finding people with good ideas and facilitating their work makes a lot more sense than trying to come up with all ideas ourselves and finding someone else to carry them out…" (Ohlbaum, 2015a). The recognition of indigenous capacities and the adaptation of knowledge to local conditions, as well as the relevance of ownership, which allows the Southern partners to make informed decisions and choices, are important characteristics of this developmental approach (Dia, 1996; UNDP, 2009). The systemic and multidimensional nature of CD focuses on the knowledge necessary to allow systems to identify their development challenges, address them, and learn from their own experience to self-renew their knowledge for the future (Alaerts & Kasperma, 2009; Morgan, 2006). Nowadays, it is clear that CD is embedded in, and cannot be isolated from, existing social systems (people, organizations, institutional settings, culture, values, politics, and power relations) that stem from historical developments.

3. CAPACITY DEVELOPMENT INTERVENTIONS AND MODALITIES

Arguably, capacity develops anyway, whether supported by capacity development interventions or not, albeit often not in the required areas or depths of expertise nor at the required speed. CD is typically understood as the process by which donors or implementing agencies foster existing capacity via CD projects, which are conducted through so-called interventions. These interventions consist of deliberate actions (i.e., activities and modalities) undertaken to enhance over time the capacity of individuals, their organizations, and societies in an effort to foster their knowledge and skills base and strengthen their own existing competencies and capabilities (Kuhl, 2009; Ubels, Acquaye-Baddoo, & Fowler, 2010). Therefore, the main objective of CD interventions should be understood as targeting knowledge, skills, and capacity in a broad sense (UNDP, 2009). In principle, capacity development aims at solving (changing) problems and improving performance (OECD, 2006), and CD interventions are planned, consciously or unconsciously, on the basis of assumptions, behavioral change, development theories, and experience. This combination of inputs leads to planned—and, more importantly, emergent (non-planned)—changes that cannot be easily measured. [1]

In practice, CD projects have predetermined objectives and indicators of achievement to facilitate, at the end of the project, the evaluation of the results achieved by the interventions. Examples of these indicators are the number of people trained, PhD or MSc/MA graduations, the number of educational curricula revised, the number of children completing primary education, and the reduction of maternal deaths during childbirth, among those at the individual level; and new or changed policies, practices, and working routines at the organizational level. In other words, projects are evaluated based on predetermined results through the use of quantified and aggregated indicators. This is part of what is known as "the results agenda," the quantification and aggregation of results to show donors and taxpayers "value for money" (Eyben, 2011; Green, 2010).

We argue that there is a need to include in the evaluation the relevance of those emergent, non-planned, changes that

occur due to—or parallel to—CD interventions and that over time strengthen and build the recipients' capacity to learn, affecting significantly the ongoing patterns of interaction and supporting social transformation.

While the process of strengthening capacity—capacity development—is nowadays conceived as the inherent responsibility of people, organizations, and/or societies themselves (OECD, 2006), external parties (mostly NGOs and developmental agencies, and primarily through consultants, trainers, and advisers) play an important supporting role in this process. This explains why the main centers of knowledge building and dissemination regarding CD are international organizations, donors, consultants, and NGOs, and not the traditional centers of knowledge such as universities or research institutes.

CD should not be something a donor does for the recipient, but rather a process undertaken jointly with the recipient, in knowledge-based collaborative partnerships (Ohlbaum, 2015a, 2015b; Santos, 2015). The essence of the discussion over CD should include the ways in which these interventions facilitate, or disrupt, livelihood improvements and social transformation.

CD interventions could be characterized as learning processes (Pearson, 2011), as they aim to strengthen or improve ways of thinking and doing (Mvulirwenande, Wehn, & Alaerts, 2015). The effect of CD interventions expands from the individuals to their organizations through the knowledge, skills, and attitudes that are developed and enhanced by the CD interventions and modalities.

It is assumed that individual learning is translated to organizational learning through procedures, routines, knowledge management, and incentive systems (UNDP, 2009; Wehn, 2014b). The enabling environment benefits from CD interventions that strengthen the policies, strategies, and legal and regulatory frameworks within which organizations can then function better. Civil society—seen as a non-state, non-market social sphere and represented by Civil Society Organizations (CSOs)—is also affected by all this learning, and at the same time, it also influences the learning processes with its culture, ambitions, and values.

In order to achieve the main goals of CD interventions—improved livelihoods and social transformation through the development of the knowledge, skills, and attitudes of individuals, procedures, and routines of organizations, interactions of sector organizations, and/or policies and frameworks of the enabling environment, as well as the strengthening of the existing ones—a large toolkit of activities or modalities can be adopted.

These various modalities, or activities, used in interventions are not implemented in isolation. They are complementary to each other and seek to strengthen capacity on different levels. One of the most traditional mechanisms for knowledge transfer is training and education. Training requires considerable amounts of finances and resources, which on many occasions are not reflected in its benefits (Blume et al., 2010). The easier quantification of training and education results (e.g., number of graduates, number of courses, or number of participants completing training) may explain why this modality is commonly used in CD interventions. However, before adopting any modality, we should ask whether it will contribute to facilitating changes in behaviors (habits and practices) or to the deep pattern of conversation or interaction toward social transformation.

Therefore, the inclusion of training in CD projects has been subject to considerable criticism. On the one hand, the application of the learning resulting from the training to the actual job is what leads to changes in work performance and therefore to the training's effectiveness. On the other hand, many authors argue that in practice, the new knowledge and skills transferred to recipients through (short-term) training is not sufficient for training to be considered effective, as there is a large gap between learning efforts and the degree to which training contents are applied to practice (Aguinis & Kraiger, 2009; Grossman & Salas, 2011). Some authors even argue that most training contents are not transferred to the job (Grossman & Salas, 2011), while others argue that only about 10% of training expenditures are transferred to the job (Georgenson, 1982).

The question of *how* adults learn and acquire relevant knowledge, skills, and competencies for their professional activities has been researched since the 1960s by Allen Tough's adult learning projects (e.g., Tough, 1967, 1968, 1971).[2] In his work, Tough finds that most adult learning occurs at the workplace and is self-directed. Eichinger and Lombardo (1996), from the Center for Creative Leadership (CCL), refer to the work done by McCall, Eichinger, and Lombardo based on 30 years of research on how executives learn to lead. The results of the study by Eichinger and Lombardo (1996) show that most learning takes place during action (i.e., learning by doing, or on-the-job learning) on the basis of what is called the "70:20:10 rule" (Jennings & Wargnier, 2015).[3]

The so-called 70:20:10 rule indicates that only 10% of relevant knowledge and expertise (for professionals) is acquired through formal training and education (e.g., courses and reading), 20% through coaching and mentoring (e.g., from people, mostly the boss), and 70% via on-the-job learning, learning by doing, and other actual experience-building activities (Eichinger & Lombardo, 1996; Jennings & Wargnier, 2015).[4] This aspect is also stressed by Nonaka and Takeuchi (1995) in their theory of knowledge creation and learning (in the corporate sector), which argues that learning takes place via a range of mostly informal interaction processes for knowledge sharing and knowledge creation (Nonaka & Takeuchi, 1995). Table 2 is taken from Jennings and Wargnier (2015) and illustrates practical examples of the 70:20:10 rule that relate to some relevant modalities commonly adopted in CD interventions.

The modalities to be implemented in each intervention are chosen in order to satisfy different learning objectives and consequently result in different learning outcomes (Preskill & Boyle, 2008). The selection of these modalities is mostly based on the experience or expertise of the implementers. However, the selection also takes into consideration the recipients' characteristics, culture, organizational resources, existing absorptive capacity, and other relevant factors (Preskill & Boyle, 2008). Consequently, there are modalities that focus on strengthening knowledge about "how" and others that strengthen knowledge about the relevant contents ("what"). While all modalities serve to strengthen individual skills, only the more interactive modalities (such as meetings, communities of practice, and action learning) cover the full spectrum of learning outcomes, not least of which are co-created meaning and understanding and stronger relationships. These are important since they serve to sustain the knowledge creation process.

Table 3 classifies common modalities adopted in CD interventions, targeted mostly at the individual level, according to their focus and their consequent different learning outputs. The table indicates some commonly used modalities (or activities) useful in addressing the co-creation of skills, meaning, and understanding, as well as others that, due to the dynamics required by the modalities, contribute to building stronger

Table 2. *Practical examples of the 70:20:10 rule*

70 Learn and develop through experience	20 Learn and develop through others	10 Learn and develop through structured courses
Experiential Learning and Development		
- Apply new learning in real situations - Cross-functional introductions, site/customer visits - Use feedback to try a new approach to an old problem - New work and solving problems within role - Work with consultants or internal experts - Increased decision-making - Champion and/or manage changes - Cover for others on leave - Interviewing - Exposure to other departments/roles - Take part in project or working group - Coordinated role swaps - Leadership activities - Stretch assignments - Project reviews - Research and apply best practice - Community activities and volunteering - Interaction with senior management (e.g., meetings, presentations) - Day-to-day research, web browsing	- Informal feedback and work debriefs - Seeking advice, asking opinions, sounding out ideas - Coaching from managers/others 360° feedback - Assessment with feedback - Structured mentoring and coaching - Learning through teams/networks - External networks/contacts - Professional association involvement or active memberships - Facilitated group discussion (e.g., action learning)	- Courses, workshops, seminars - eLearning - Professional qualifications - Accreditation - Certification - Formal education (e.g., university)

Source: From Jennings and Wargnier (2015, pp. 15–16).

Table 3. *Strengthening individual capacity—capacity development modalities and their focus*

Categorization of modalities (applying the 70:20:10 rule)	Modality	Knowing how	Knowing what	Co-creation of meaning and understanding	Building stronger relationships	Skill development
20	Internships	✓				✓
10	Written materials		✓			✓
70	Meetings	✓	✓	✓	✓	✓
20	Appreciative inquiry			✓		✓
20	eLearning	✓		✓		✓
20	Coaching or mentoring	✓		✓	✓	✓
10	Training	✓	✓	✓		✓
70	Communities of practice	✓	✓	✓	✓	✓
20	Action learning (in peer group)	✓	✓	✓	✓	

Source: Developed by the authors based on Preskill and Boyle (2008) and complemented based on Jennings and Wargnier (2015).

relationships among participants and stakeholders. In addition, Table 3 structures these modalities according to the 70:20:10 rule in order to complement their understanding.

When capacity development objectives extend beyond strengthening individual skills, competencies, and expertise to address organizational and sector performance, the modalities adopted draw increasingly on organizational design (e.g., structure as well as incentive systems), change management, and knowledge management (i.e., processes).

While it may be true that learning "takes place inside individual human heads, [and contributes to organizational learning] in two ways: (a) by the learning of its members, or (b) by ingesting new members who have knowledge the organization did not previously have" (Simon, 1996, p. 176), it would be a mistake to conclude that organizational learning is nothing but the cumulative result of an organization's members (Hedberg, 1981). For example, an organization may suffer from low performance because its structure does not support change in habits and practices, even though the individuals that make up the organization possess relevant and up-to-date knowledge. In such a situation, it is necessary to use CD modalities targeting cultural change, which involve interventions related to fostering team work, responsibility sharing, partnerships, and strategic planning, among others (Levinger, 2015).

Figure 1 illustrates the salience of process consultation as a CD modality as well as the strengthening of interactions and relationships (between individuals and organizations). Based on Denker and Lutz (2014) and DFID (2003), Figure 1 adds to their conceptualization a focus on modalities for leadership development. This is based both on the acknowledged role of leaders and change agents for implementing change as well as the realization from transformational leadership theory that leadership can indeed be fostered and is not merely a talent that some individuals have and others do not (Burns, 1978; Lincklaen Arriens & Wehn de Montalvo, 2013). Similarly, action learning has been added to the examples, since it provides a salient CD modality for enabling joint, job-related learning within groups of peers (from the same or similar organizations) to strengthen both the expertise of the involved individuals as well as organizational procedures; as such, it is also highly relevant for strengthening organizational processes.

4. EVALUATION OF CAPACITY DEVELOPMENT

Evaluation is the most formal way to examine work and generate feedback (Britton, 2010). Project evaluation is a common way to account for the investments provided by project funders and the efforts of relevant stakeholders (Carman & Fredericks, 2010; Preskill & Boyle, 2008). As capacity development projects have increased, the evaluation of CD project outcomes has taken on increasing relevance. The increasing pressure to deliver measurable CD results, or the so-called "results agenda," is pushing the CD donor community to increasingly focus on projects with tangible (and mostly pre-defined) results.

The value for money debate is so focused on providing quantified aggregated numbers (e.g., number of girls completing primary education, number of children receiving breakfast at school, number of graduates) for the financed CD interventions, that the focus on processes and non-planned changes and their effects on social transformation is being pushed aside.

There are two main schools of thought regarding the monitoring and evaluation of capacity development, referred to by
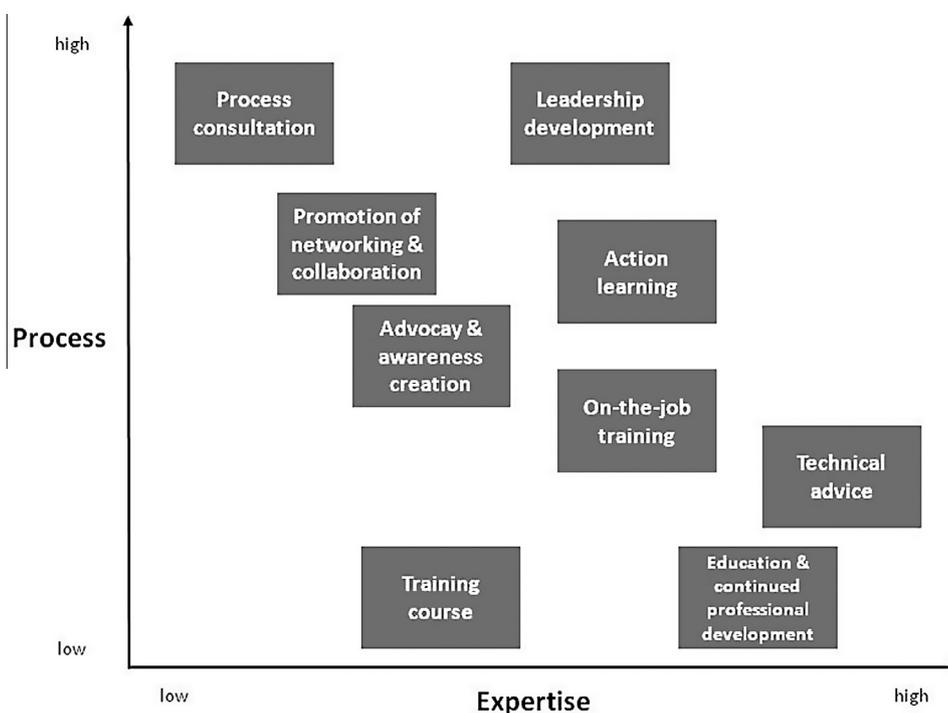
Figure 1. *Examples of modalities for strengthening organizational capacity, by level of fostered expertise and process. Source: Based on Denker and Lutz (2014) and DFID (2003).*

Watson (2010) as technocratic thinking (called the linear connection system by other authors) and the complex adaptive system (CAS) approach (Watson, 2010). Although both lines of evaluation have their own benefits and uses, their application depends on the circumstances and the reasons behind the CD intervention in question (Baser & Morgan, 2008; Morgan, Land, & Baser, 2005; Watson, 2010).

The first approach is mostly used by international donors to evaluate the effectiveness of their large amounts of funding in light of their own accountability needs (Morgan et al., 2005; Watson, 2010). Under this approach, international organizations recognize three main points of entrance to CD (UNDP, 2009). The first entry point is National Development Strategies (NDS), which are defined by national governments. These consist of broader areas in which assessments of needs and capacities are required and then used as a basis to define CD strategies. The second entry point for CD targets those sectors on which donors and local governments have placed their priorities. Nowadays, food security, water, education, environment, and security are among the sectors included in most donors' strategic plans. The third point of entry is that presented by themes in the international agenda. Issues such as climate change, disaster risk reduction and resilience, national implementation capacities, and gender empowerment are among those most commonly addressed by the international agenda and consequently by donors (UNDP, 2009).

This approach is commonly characterized by the use of logical frameworks (*logframe*) for program planning and monitoring (Watson, 2010). The logical (or project) framework specifies indicators that are related to assessing the progress of the project toward achieving its objectives. However, in practice, due to the time and resource constraints on the implementers, these indicators often become the objective itself and not a proxy for evaluating the CD goal set forth for the intervention.

As Eyben (2011), currently from the Institute of Development Studies and formerly by the Chief Social Development Adviser at DfID, states: "[S]ometimes...[a] bounded-problem-approach to change, as typified by the logical framework, is going to work....[However, although] the major concern is about the institutional and financial sustainability once the intervention ends, these [aspects] are not addressed as an integral part of the CD intervention design." Eyben continues with an example of how an intervention targeting the global eradication of poliomyelitis was a success in all but four countries, and how, 2 years later, the disease was back in 19 additional countries. The issue, she explains, was that the intervention did not pay enough attention to the changes needed to sustain these results, because the drive preference was given to the achievement of the pre-defined results (Eyben, 2011).

Under the technocratic approach, an increase in (organizational) performance seems to be considered as a proxy for capacity (Morgan et al., 2005; Watson, 2010). The rationale behind this logic is that there is a linear connection between the provision of inputs and the delivery of (previously defined) outputs (i.e., cause–effect relationship), which (under certain assumptions, also stated in the *logframe*) leads to an improvement in performance and the achievement of the development goals established by the CD project (Watson, 2010). Examples of this type of approach are the approaches known as "managing for development results," "results-based management," and "results-based approaches" (UNDP, 2010).

The four key components of the results-based approach are: strategic planning, impact, outcomes, and outputs (UNDP, 2010). The strategic planning process seeks to identify the impact, outcome, and outputs, as well as the definition of the development goals targeted in the CD project. To achieve these development goals, the strategic plan should identify the specific changes (or outcomes) that most occur within various

systems and then coordinate them with the stakeholders of the project (UNDP, 2010).

Impact is related to the actual or intended change in human development, measured by people's well-being, which can be defined by the National Development Goals (NDGs) targeted by the project (UNDP, 2010). Measuring the impact of CD interventions using technical performance indicators as a proxy has been criticized as a misleading indication of the actual impact of CD efforts, due to the time involved for capacity changes to translate into performance improvements (Wehn de Montalvo & Alaerts, 2013).

An outcome is the actual or intended change in development condition that interventions are seeking to support. It is measured by the change in institutional performance, stability, and adaptability; these three elements influence national institutions, which then influence the NDGs addressed by the impact (UNDP, 2010). Mvulirwenande, Wehn, and Alaerts (2014) and Pascual Sanz et al. (2013) argue that performance improvements and capacity improvements are different—but nevertheless related—outcomes of CD interventions and, therefore, should both serve as sources of evidence about the effectiveness of CD interventions.

Outputs are short-term development results produced by project and non-project activities (UNDP, 2010). They are defined as products produced or services provided based on capacity development's core issues (i.e., institutional arrangements, leadership, knowledge, and accountability).

The second main type of evaluation approach, complex adaptive systems approaches (CAS), focuses on capturing changes in the behavior of and relationships between the direct participants of the CD process, rather than specified outcomes in terms of traditional technical performance indicators. It takes into account the complex nature of capacity and capacity development and is grounded in participatory approaches. CAS approaches are also open to recognizing intended and unintended changes resulting from CD (Pascual Sanz et al., 2013).

CAS is mostly adopted by NGOs and developmental agencies to emphasize learning from experience as a means to improve capacity (Morgan et al., 2005; Watson, 2010). It encourages interaction between stakeholders in the selection, management, and evaluation of specific CD modalities in complex settings—from the outset of interventions—by discussion and reflection about the assumptions underlying the envisaged change and the conditions that might be needed for change to emerge (taking into account power, culture, systems, and actors), from multiple levels and with different perspectives (Datta, Shaxson, & Pellini, 2012; Ortiz Aragon & Giles Macedo, 2010). Anecdotes based on experience are used to build knowledge and awareness that changes are needed and that the development and improvement of capacity is associated with multiple elements, not all of which are easy to plan (Morgan et al., 2005; Watson, 2010). Specific examples are the Most Significant Change approach (Dart & Davies, 2003), which involves the collection of stories of change emanating from the field level, Theory of Change (Stachowiak, 2009; Vogel, 2012), and Outcome Mapping or OM (Earl, Carden, Patton, & Smutylo, 2001), which takes a learning-based and use-driven view of CD.

According to Jones and Hearn (2009), the key principles underlying OM are:

(i) Defining project aims and indicators of success in terms of changes in behavior, interactions, mind-sets, and motivations of actors;

(ii) continuous learning and flexibility via cyclical, iterative, and reflexive management activities;
(iii) involvement of stakeholders and their reflection on relationships and responsibilities, and two-way accountability;
(iv) non-linearity and contribution as opposed to attribution and control via collective ownership of processes of social change.

The most recent evaluation methodology in the CAS category is based on the European Centre for Development Policy Management's (ECDPM) five capabilities framework (Baser & Morgan, 2008; Engel, Keijzer, & Land, 2007; Keijzer, Spierings, Phlix, & Fowler, 2011). This proposes to assess change in five core capabilities that affect the capacity and performance of the overall organization or system, namely the capability (1) to survive and act, (2) to generate development results, (3) to relate (in order to achieve objectives), (4) to adapt and self-renew, and, connecting the previous capabilities, (5) to achieve coherence (Engel et al., 2007).

The CAS approaches present a fundamental shift in thinking about CD goals, the selection of CD modalities, and their attribution to change as opposed to the distinct outputs, outcomes, and impacts of the results-based approach.

In projects in which local governments and organizations are also contributing financially to CD, adopting CAS is recommendable (Watson, 2010). CAS-type evaluation methods give greater accountability to domestic stakeholders and service users, as they encourage higher levels of ownership in the CD process and promote higher strategic thinking by the beneficiary partners than most technocratic approaches (Ortiz & Taylor, 2009; Watson, 2010). Nevertheless, this approach can be criticized for leaning toward the other extreme, refusing to define tight indicators, and using such broad conceptualizations of (collective) capacity that improvements thereof are literally pre-defined.

A hybrid approach combining elements from the results-based and the system approaches can provide a practical middle ground, allowing insights into the effects of interactions between involved stakeholders and expected as well as unexpected changes while also making some efforts to predict (and manage for) tangible outcomes.

5. EVALUATION METHODS

It is clear that donor organizations want to know the benefits of their financial investments in capacity development, particularly at times when finances are scarce. Evaluating the direct relationship between capacity development and financial investments, as well as whether the knowledge and skills acquired in specific training are indeed used at work, is not an easy task (Blume et al., 2010). Ex-post sustainability has become a frequent item in the international discussion over CD interventions in the South and a clear concern of most donors (Clark et al., 2003). The absence of a common method for measuring the effectiveness of CD interventions and the severe limitations on quantifying the relationship between investment and output complicates the quantification of their benefits (Pascual Sanz et al., 2013).

The discrepancy between the investments undertaken in CD and their low effects on the development of the recipient countries has raised criticism, not only among external agents, but also among development organizations (Berg, 1993; Kuhl, 2009; Morgan & Baser, 1993). In providing a quantitative measurement for the impact, outcome, and outputs achieved

by CD initiatives, international organizations have contributed to the creation of indicators and the application of several evaluation methods. International organizations, such as the European Commission, have produced toolbox documents with overviews of evaluation concepts and methodologies (Grossman & Salas, 2011).

Examples of these measurements include the attempts to calculate economic rates of return for capacity development in irrigation reform programs by the World Bank (Darghouth, 2007; Gonzales & Salman, 2002) and the OECD efforts to assess the economic returns of secondary and advanced education. These types of efforts are time-consuming and require expertise and financial resources that are not available to most implementing CD organizations.

Therefore, to evaluate the results of their projects, many CD donors base their evaluation methods on the detailed project plan (which in fact uses a *logframe* or similar matrix to set up, monitor, and evaluate the project). Others use economic theoretical foundations, allowing them to measure in a qualitative or quantitative way the output of their projects. Relevant assessment or evaluation methods commonly used in assessing CD projects and interventions are: case-studies, cost–benefit and SWOP analysis, control group approaches, productivity studies, and macro and micro methods.

Case studies. This type of methodology is used in the assessment of qualitative insights of the unit of study. It describes relationships among the actors of the system, and it enables the identification (through a story) of details of the inputs and outputs of the intervention. It is used in cases where organizational efficiency and the competitiveness of quality elements are evaluated (Polt & Rojo, 2002). The CAS school of evaluation makes frequent use of this type of methodology. The downside of this type of method is that results are not generalizable (although the lessons could be). However, they facilitate a clearer and more detailed understanding of how the enabling environment (context) affects and shapes impacts (Polt & Rojo, 2002).

Cost–benefit analysis (CBA) and Strengths, Weakness, Opportunities, & Problems (SWOP). The special feature of these methods is their effort to evaluate quantitatively all economic and social consequences of the project or intervention. The distribution of costs and benefits is not considered and regards the market as the environment in which costs and benefits are reflected. The methods require important assumptions and subjectivity; results are not easily comparable between cases. The objective of this methodology is to assess the efficiency of the project by maximizing the difference between benefits and costs. Three instruments are used: the net present value of the investment, the internal rate of return, and the cost–benefit ratio.

Control group approaches. This methodology captures the effect of the program on the recipients. Its implementation is widely adopted in societal research, as well as social policy interventions for poverty eradication (e.g., Abhijit, Duflo, Glennerster, & Kinnan, 2015); however, it is a costly procedure and results are subject to debate by quantitative scientists (regarding the cause–effect relationship for specific interventions) and development theorists regarding ethics (concerned about depriving the control group of potential improvements). For the sake of robustness of results, the method should be implemented with strict statistical conditions (i.e., statistically significant sample sizes, stratified samples as test population). Reliability depends entirely on the quality and size of the data collected. The methodology is commonly applied in combination with productivity analysis or benchmarking.

Productivity analysis. This allows measuring quantitatively the amount of output or added value produced by a given set of inputs. The units of analysis are the micro, meso, or macro level (Aguinis & Kraiger, 2009). The analysis can be conducted with a single production factor or several (i.e., single-factor productivity or multifactor productivity). The methodology requires data covering some period of time. The major difficulties of this technique are the assumptions required for the measurement of variables and the availability and quality of the data provided (Grossman & Salas, 2011). In the case of CD projects, this technique has been used in combination with control group approaches to assess the contribution of the strategies of projects to recipients and their organizations' productivity (for example, Kirkpatrick and Kirkpatrick (2006).

Micro and macro methodologies. Micro econometric methodologies allow the evaluation of the existence of effects and the rates of return given an external influence. Macroeconomic methods are better suited to capturing spillovers and long-term effects of external interventions on productivity and economic welfare (Polt & Rojo, 2002). Other commonly used methodologies within this category are micro and macroeconomic modeling and simulation approaches. The first ones analyze the effect of an intervention at the level of individuals or firms, while the latter estimates the broader socio-economic impact of the intervention(s) (Polt & Rojo, 2002).

The incorporation of learning in the day-to-day operation environment of the recipient is a key factor in post-project sustainability (Travisan, 2002). The quantitative measurement of the output of these types of projects or the measurement of return on investment of these types of projects is a complicated task. Most of the commonly adopted evaluation methods target socio-economic outputs as well as identified institutional changes. These methodologies provide efficient assessments of the performance of projects that by nature target specific changes and whose outputs are clearly identified. This is the case of technological development projects in industries, firms, or organizations (units of study in most academic research). However, when dealing with CD projects, there is not robust methodology for measuring in quantitative terms the impact these projects have at the levels of the individual or the society, let alone their sustainability over time.

Many studies combine several data collection instruments to develop a comprehensive picture of capacity. Consequently, many combine quantitative and qualitative data to analyze it at the levels of inputs, process, outputs, and outcomes (Brown, LaFond, & Macintyre, 2001). UNDP (2013) presents an interesting overview of innovative tools used in monitoring and evaluating CD for sustainable development.

Although most of the tools presented are used in monitoring, the paper presents insights into the constant efforts of organizations and researchers to use available techniques to collect data and information in order to evaluate quantitatively their CD efforts (UNDP, 2013). Examples of micro narratives analyzed through algorithms and specialized software (e.g., Atlas Ti, Sensemaker) as well as multi-level mixed evaluation methods and outcome harvesting provide a clear view of the use of hybrid tools combining quantitative with qualitative information for the evaluation of impacts and outputs of

CD projects (Mertens & Hesse-Biber, 2013; UNDP, 2013). Furthermore, subjective as well as more objective measures can be combined to obtain insights into the extent of capacity change at individual and organizational levels (Pascual Sanz et al., 2013), although extensive efforts are still required to collect these insights.

6. DISCUSSION

As presented in this study, over time capacity development has evolved from its original understanding as being synonymous with education and training to a system of intervention that fosters the knowledge base and the capacity of individuals and organizations by creating learning opportunities and assisting with the generation and acquisition of new knowledge and skills and aiming to facilitate systemic and lasting change. The formulation of the post-2015 development agenda and the finalization of the SDGs highlight that the emphasis placed on capacity development in the development discourse has also changed considerably, with CD now constituting a core vehicle for *sustainably* achieving the SDGs (Wehn, 2014a).

The study explains how, in practice, one CD project implements diverse CD interventions in order to tackle its specific CD objectives by initiating or facilitating change processes. These interventions can consist, for example, of formal (basic and higher) education and distance learning; advisory services and peer learning to foster change and innovation within organizations and institutions; partnerships, knowledge networks, and collaboration for social learning; and leadership development of individuals and organizations to effect change.

As resources are scarce, donor agencies and policy-makers need to know the value of CD in order to best prioritize their investments. However, due to the nature of capacity development projects (consisting of change processes by initiating change in people, organizations, and/or their enabling environment), measuring the return rate between the project cost and its impact remains a difficult task. We should not forget that CD is a complex process, as it deals with changes of habits, practices, and paradigms. The relevance of non-planned results and the partnerships with recipients play a key role in developing, or strengthening, the learning behind social transformation. However, these aspects of integrating the knowing with the doing are not easily quantified and on many occasions do not fit the pre-established indicators and quantified results.

While interventions and modalities can be related to immediate outputs (e.g., number of trained staff, number of curricula revised, participation of female staff, knowledge management procedures set up/revised, human resources management policies modified, sector interaction platforms installed, and number of actors connected, among others), the relationship between these outputs and outcomes remains hard to quantify, if not elusive, due to the non-linear relation between these two elements. The same occurs between outcomes and long-term impact. Knowledge and capacity are hard to quantify, as they develop in organizations and institutions across sectors via long-run, complex processes driven by local demand and local dynamics, which need to be thoughtfully organized and supported. Their success requires both leadership and ownership by the local counterparts. Moreover, perceptions of success are highly reliant on the definition and measures of success (in whose view, according to what measurement, and at what moment in time).

CD by its very nature is a hard-to-plan process with a large percentage of endogeneity. Linear approaches to CD have clear limitations when it comes to specifying (ex-ante and ex-post) outputs, outcomes, and impacts of capacity development interventions. Clear performance indicators are difficult to attribute to each CD project, since many projects are interrelated and are implemented at the same time as others (i.e., different projects from different donors are sometimes run simultaneously at the same location and by the same organization, even involving the same participants). Swinging to the other extreme and relying on a systems-based approach alone can lead to early project fund depletion with limited change.

The type of measurement methodology used to evaluate impact depends not only on whether we want to assess simple or complex knowledge types and the area of application (e.g., organization *versus* sector), but also on the fundamental conceptualization of CD that projects are based on to begin with, since this choice inherently determines the evaluation approach. We also need to consider that this activity is time-consuming and expensive and requires expert evaluation rather than direct objective measurement.

This is a real challenge faced by project implementers. On the one hand, the time frame for implementation and the objectives and goals of the project are already determined and agreed upon beforehand with the donors. Therefore, *logframes* and other traditional reporting and evaluation tools required by the donors are the most commonly adopted evaluation techniques. On the other hand, there is an increasing awareness and need for hybrid, multidimensional evaluation methodologies that reflect the complexity of the environment in which CD projects are implemented.

Traditional methods provide suitable tools to give quantified indications of the results of activities such as education and training. However, when time frame horizons are wider and the evaluation involves larger sectors and a wider range of activities, the measurement of outcomes and impacts becomes more complex. Alongside the systemic view of capacity arises the challenge of capturing changes to the "system" in question.

7. CONCLUSIONS

The study concludes that existing methodologies and approaches are not sufficient, and new, hybrid or multi-path approaches to capture changes in capacity and their contribution to results still require further methodological development in order to effectively provide a basis for careful decisions regarding future CD efforts and investments. The research joins the ongoing results agenda debate, by arguing that in the case of CD interventions, we should not forget that we are dealing with projects that, by their nature, rely more on non-planned changes than on the pre-defined indicators and results to contribute to patterns of change toward livelihood improvements and social transformation. Therefore, the research concludes that current results-based evaluation methods may provide a good measurement of CD outputs and outcomes that can be captured by indicators to report progress to donors. However, they are not able to measure impact, as CD projects, restricted by previously agreed budgets, resources, and time frames, are usually not designed to evaluate the sustainability of change and its impact over the medium or long term.

NOTES

1. Emergent changes are defined as actions, adaptations, and alterations that produce fundamental change without an *a priori* intention to do so (Lifvergen, Docherty, & Shain, 2011; van Kemenade, 2014).

2. More on Tough's work can be found at http://allentough.com.

3. Although Tough at the time did not refer to the 70:20:10 rule, he later acknowledged that he also found this relationship in his research (70:20:10 Forum Pty Ltd, 2015).

4. This framework for learning nowadays characterizes approaches to learning in many organizations, including commercial ones. For example, in 2002, Charles Jennings, Reuter's Chief Learning Officer, developed a strategic model for practical implementation at Reuters extending the 70:20:10 rule into strategic tools for the use of human resources, talent, and learning for professionals.

REFERENCES

Abhijit, B., Duflo, E., Glennerster, R., & Kinnan, C. (2015). The miracle of microfinance? Evidence from a randomized evaluation. *American Economic Journal: Applied Economics, 7*(1), 22–53. http://dx.doi.org/10.1257/app.20130533.

Aguinis, H., & Kraiger, K. (2009). Benefits of training and development for individuals and teams, organizations and society. *Annual Review of Psychology, 60*, 451–474. http://dx.doi.org/10.1146/annurev.psych.60.110707.163505.

Alaerts, G., & Kasperma, J. (2009). Progress and challenges in knowledge and capacity development. In M. Blokland, G. Alaerts, & J. Kasperma (Eds.), *Capacity development for improved water management*. London, UK: Taylor and Francis.

Alley, K., & Negretto, G. (1999). *Literature review: Definitions of capacity building and implications for monitoring and evaluation*. New York: UNICEF, Division of Evaluation, Policy and Planning.

Argyris, C., & Schön, D. (1978). *Organizational learning: A theory of action perspective*. MA: Addison-Wesley.

Banks, N., Hulme, D., & Edwards, M. (2015). NGOs, states, and donors revisited: Still too close for comfort?. *World Development, 66*, 707–718. http://dx.doi.org/10.1016/j.worlddev.2014.09.028.

Baser, H., & Morgan, P. (2008). Capacity, change and performance. In ECDPM (Ed.), *Discussion paper (59B)*. Maastricht, NL: European Center for Development Policy Management (ECDPM).

Berg, E. J. (1993). *Rethinking technical cooperation reforms for capacity building in Africa*. New York: UNDP.

Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. *Journal of Management, 36*(4), 1040–1064. http://dx.doi.org/10.1177/0149206309352880.

Britton, B. (2010). Self-reflection. Monitoring and evaluation for personal learning. In J. Ubels, N. A. Acquaye-Baddoo, & A. Fowler (Eds.), *Capacity development in practice* (pp. 264–276). London: Earthscan.

Brown, L., LaFond, A., & Macintyre, K. (2001). *Measuring capacity building*. Chapter Hill: MEASURE Evaluation Project.

Burns, J. (1978). *Leadership*. New York: Harper & Row Publishers.

Carman, J. G., & Fredericks, K. A. (2010). Evaluation capacity and nonprofit organizations: Is the glass half-empty or half-full?. *American Journal of Evaluation, 31*(1), 84–104. http://dx.doi.org/10.1177/1098214009352361.

Clark, N., Hall, A., Sulaiman, R., & Naik, G. (2003). Research as capacity building: The case of an NGO facilitated post-harvesting innovation system for the Himalayan Hills. *World Development, 31*(11), 1845–1863. http://dx.doi.org/10.1016/j.worlddev.2003.04.001.

Cohen, J. M. (1994). *Capacity building, methodology and the language of discourse in development studies*. Cambridge, Mass.: Harvard Institute for International Development.

Darghouth, S. (2007). *Emerging public–private partnerships in irrigation development and management*. Washington, DC: World Bank.

Dart, J., & Davies, R. (2003). A dialogical, story-based evaluation tool: The most significant change technique. *American Journal of Evaluation, 24*(2), 137–155. http://dx.doi.org/10.1177/109821400302400202.

Datta, A., Shaxson, L., & Pellini, A. (2012). *Capacity, complexity and consulting lessons from managing capacity development projects*. London: Overseas Development Institute.

Denker, C., & Lutz, T. (2014). *Handbook operationalization of capacity development in the Uganda water and environment sector*. Kampala, Uganda: Ministry of Water and Environment.

DFID (2003). *Promoting institutional and organisational development, department for international development issues*. London: DFID.

Dia, M. (1996). *Africa's management in the 1990s and beyond*. Washington, DC: The World Bank.

Earl, S., Carden, F., Patton, M. Q., & Smutylo, T. (2001). *Outcome mapping: Building, learning and reflection into development programs*. Ottawa: International Development Research Center (IDRC).

Edquist, C. (Ed.) (1997). *Systems of innovation approaches: Technologies, institutions and organizations*. London: Pinter, Casell Academic.

Eichinger, R., & Lombardo, M. (1996). *The career architect development planner*. Minneapolis: Lominger Limited.

Enemark, S., & Ahene, R. (2002). *Capacity building in land management – Implementing land policy reforms in Malawi*. Paper presented at the FIG XXII International Congress. Washington, DC.

Engel, P., Keijzer, N., & Land, T. (2007). *A balanced approach to monitoring and evaluating capacity and performance. A proposal for a framework*. Discussion paper. Maastricht: ECDPM.

European Commission JRC-IPTS & Joanneum Research (2002). RTD evaluation toolbox – Assessing the socio-economic impact of RTD-policies. In G. Fahrenkrog, W. Polt, J. Rojo, A. Tubke, & K. Zinocker (Eds.), *IPTS technical report series*. Sevilla, Spain: European Commission Joint Research Centre and Institute for Prospective Technological Studies.

Eyben, R. (2011). 'Stuff happens': The risks of a results agenda. Retrieved from <http://oxfamblogs.org/fp2p/stuff-happens-the-risks-of-a-results-agenda-guest-post-from-rosalind-eyben/>.

Forum Pty Ltd. (2015). The 70:20:10 framework. What is it? Retrieved January 18, 2015, from <https://www.702010forum.com/about-702010-framework>.

Fowler, A. (1997). *Striking a balance: A guide to enhancing the effectiveness of non-governmental organizations in international development*. London: Earthscan.

Georgenson, D. L. (1982). The problem of transfer calls for partnership. *Training and Development Journal, 36*(10), 75–78.

Godfrey, M., Sophal, C., Kato, T., Piseth, L. V., Dorina, P., Saravy, T. S., et al. (2002). Technical assistance and capacity development in an aid-dependent economy: The experience of Cambodia. *World Development, 30*(3), 355–373.

Gonzales, F. J., & Salman, S. M. A. (2002). Institutional reform for irrigation and drainage: Proceedings of a World Bank workshop. In *World Bank technical paper (524)*. Washington, DC: World Bank.

Green, D. (2010). Is the aid industry's audit culture becoming a threat to accountability? Retrieved from <http://oxfamblogs.org/fp2p/is-the-aid-industrys-audit-culture-becoming-a-threat-to-accountability/>.

Green, D. (2015). What do we know about the politics of reducing inequality? Not much – It's time to find out. Blog. Retrieved from <http://oxfamblogs.org/fp2p/what-do-we-know-about-the-politics-of-reducing-inequality-not-much-its-time-to-find-out/>.

Grossman, R., & Salas, E. (2011). The transfer of training: What really matters?. *International Journal of Training and Development, 15*(2), 103–120.

Guicquero, Y. (2015). Financing sustainable development: The resources are already in the south. *Global Views. #Fin4Dev*. Retrieved July 2, 2015, from <https://www.devex.com/news/financing-sustainable-development-the-resources-are-already-in-the-south-86410>.

Gunzburg, T. (2015). Time for the EU to lay its cards on the table. *Global Views. #FIN4DEV*. Retrieved June 30, 2015, from <https://www.devex.com/news/time-for-the-eu-to-lay-its-cards-on-the-table-86404>.

Hedberg, B. (1981). How organizations learn and unlearn. In P. Nystrom, & W. H. Starbuck (Eds.), *Handbook of organizational design*. New York: Oxford University Press.

Hencke, D. (2015). IDC chair candidates promise tougher DfID scrutiny, value for money. *Inside Development. The Future of DFID*. Retrieved June 5, 2015, from <https://www.devex.com/news/idc-chair-candi-dates-promise-tougher-dfid-scrutiny-value-for-money-86360>.

HLF3. (2008). *Accra agenda for action*. Paper presented at the Third High Level Forum on Aid Effectiveness, Accra, Ghana.

Jennings, C., & Wargnier, J. (2015). *Effective learning with 70:20:10. The new frontier for the extended enterprise*. CrossKnowledge.

Jones, H., & Hearn, S. (2009). *Outcome mapping: A realistic alternative for planning, monitoring and evaluation*. London: ODI.

Joshi, D. K., Hughes, B. B., & Sisk, T. D. (2015). Improving governance for the post-2015 sustainable development goals: Scenario forecasting the next 50 years. *World Development, 70*, 286–302. http://dx.doi.org/10.1016/j.worlddev.2015.01.013.

Jozwiak, G. (2015). The future of DfID – And UK development cooperation. *Inside Development. UK Aid*. Retrieved from Devex. The Development Newswire website: <https://www.devex.com/news/the-future-of-dfid-and-uk-development-cooperation-85455>.

Kaplan, A. (2000). Capacity building: Shifting the paradigms of practice. *Development in Practice, 10*(3 & 4), 517–526.

Keijzer, N., Spierings, E., Phlix, G., & Fowler, A. (2011). *Bringing the invisible into perspective*. Reference paper for using the 5Cs framework to plan, monitor and evaluate capacity and results of capacity development processes. Maastricht: ECDPM.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels* (3rd ed.). San Francisco, CA: Berrett-Koehler.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

Kuhl, S. (2009). Capacity development as the model for development aid organizations. *Development and Change, 40*(3), 551–577.

Levinger, B. (2015). Capacity building. More than just training. In K. Warren (Ed.), *Inside development*. <https://www.devex.com/news/ca-pacity-building-more-than-just-training-85376>: Devex.com.

Lifvergen, S., Docherty, P., & Shain, A. B. R. (2011). Toward a sustainable healthcare system: Transformation through participation. In S. A. Mohrman, A. B. R. Shani, & C. Worley (Eds.), *Organizing for sustainable effectiveness*. London: Emerald Books.

Lincklaen Arriens, W. T., & Wehn de Montalvo, U. (2013). Exploring water leadership. *Water Policy, 15*, 15–41. http://dx.doi.org/10.2166/wp.2013.010.

Lusthaus, C., Adrien, M.-H., & Perstinger, M. (1999). *Capacity development: Definitions, issues and implications for planning, monitoring and evaluation*. Universalia occasional papers. Montreal: Universalia.

Mansell, R., & Wehn, U. (Eds.) (1998). *Knowledge societies – information technology for sustainable development*. Oxford: Oxford University Press.

Mertens, D., & Hesse-Biber, S. (2013). Mixed methods and credibility of evidence in evaluation. *New Directions for Evaluation, 138*, 5–13. http://dx.doi.org/10.1002/ev.20053.

Morgan, G. (1989). *Creative organizational theory: A resource book*. Newbury Park, CA: SAGE Publications.

Morgan, P. (1998). *Capacity and capacity development – some strategies*. Note prepared for the Political and Social Policies Division. Ottawa: CIDA.

Morgan, P. (1999). *An update on the performance monitoring of capacity development programs: What are we learning?*. Ottawa: CIDA.

Morgan, P. (2006). *The concept of capacity*. Maastricht: European Centre for Development Policy Management (ECDPM).

Morgan, P., & Baser, H. (1993). *Making technical cooperation more effective. New approaches in international development*. Hull: CIDA, Technical Cooperation Division.

Morgan, P., Land, T., & Baser, H. (2005). *Study on capacity, change and performance*. Discussion paper. Maastricht, NL: European Centre for Development Policy Management (ECDPM).

Mvulirwenande, S., Wehn, U., & Alaerts, G. (2014). *Evaluating capacity development in the water sector: Challenges and progress*. Paper presented at the 37th WEDC international conference "Knowledge management in water utilities: From challenges to priorities." Hanoi, Vietnam.

Mvulirwenande, S., Wehn, U., & Alaerts, G. (2015). *Closing the learning cycle in capacity development. A two-step approach to impact evaluation*. Paper presented at the 38th WEDC international conference. Water, sanitation and hygiene services beyond 2015: Improving access and sustainability. Loughborough University.

Nelson, R., & Winter, S. (1982). *An evolutionary theory of economic change*. Harvard University Press.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge creating company. How Japanese companies create the dynamics of innovation*. New York: Oxford University Press.

OECD. (1996). *Shaping the 21st century: The context of development co-operation*. Paris: OECD.

OECD. (2006). The challenge of capacity development. Working towards good practice. In *DAC guidelines and reference series*. Paris: OECD.

OECD. (2007–2012). Country programmable aid (CPA) (Aid Architecture). Available from OECD.org Interactive CPA Charts. Retrieved October 17, 2014 <http://www.oecd.org/development/aid-architecture/cpa.htm>.

OECD. (2011). *Busan partnership for effective development co-operation*. Paper presented at the Fourth High Level Forum on aid effectiveness. Busan, Korea.

Ohlbaum, D. (2015a). The illusion of control. *Foreign aid effectiveness: A radical rethink*. Retrieved February 19, 2015, from <https://www.devex.com/news/the-illusion-of-control-85523>.

Ohlbaum, D. (2015b). The siren song of technical assistance. *Global Views. Funding trends. Foreign aid effectiveness. A radical rethink (3/7)*. Retrieved February 25, 2015, from <https://www.devex.com/news/the-siren-song-of-technical-assistance-85539>.

Ortiz Aragon, A., & Giles Macedo, J. C. (2010). A 'systemic theories of change' approach for purposeful capacity development. *IDS Bulletin, 41*(3), 87–99.

Ortiz, A., & Taylor, P. (2009). Learning purposefully in capacity development: Why, what and when to measure?. In International Institute for Education and Planning (IIEP) (Ed.), *Rethinking capacity development*. Paris: UNESCO-IIEP.

Pascual Sanz, M., Veenstra, S., Wehn de Montalvo, U., van Tulder, R., & Alaerts, G. (2013). What counts as "results" in capacity development partnerships between water operators? A multi-path approach toward accountability, adaptation and learning. *Water Policy, 15*, 242–266. http://dx.doi.org/10.2166/wp.2013.022.

Pearson, J. (2011). Training and beyond: Seeking better practices for capacity development. *OECD development cooperation working papers, 1*.

Polt, W., & Rojo, J. (2002). Evaluation methodologies. In G. Fahrenkrog, W. Polt, J. Rojo, A. Tubke, & K. Zinocker (Eds.), *RTD-evaluation toolbox assessing the socio-economic impact of RTD-policies*. Sevilla, SP: European Commission, JRC-IPTS and Joanneum Research.

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation, 29*(4), 443–459. http://dx.doi.org/10.1177/1098214008324182.

Raynard, P. (2000). *Mapping accountability in humanitarian assistance*. Report to Biannual Meeting. UK: Active Learning Network for Accountability and Performance in Humanitarian Assistance (AL-NAP).

Roberts, D. (2013). Managing knowledge results: An exploration of the systems and practices of development agencies. In *Capacity development and results*. World Bank Institute.

Santos, L. A. (2015). What makes knowledge partnerships for development work. *Devex Impact. Inside Development. Knowledge Partnerships*. Retrieved June 5, 2015, from <https://www.devex.com/news/what-makes-knowledge-partnerships-for-development-work-86206>.

Sastre Merino, S., & de los Rios Carmenado, I. (2012). Capacity building in developing countries. *Procedia – Social and Behavioral Sciences, 46*, 960–967. http://dx.doi.org/10.1016/j.sbspro.2012.05.231.

Senge, P. M. (1990). *The fifth discipline. The art and practice of the learning organization*. Bantam Doubleday Dell Publishing Group, Inc..

Simon, H. (1996). Bounded rationality and organizational learning. In W. M. Cohen, & L. S. Sproull (Eds.), *Organizational learning*. London: Sage Publications.

Smillie, I. (1995). *The alms bazaar: Altruism under fire – Non-profit organizations and international development*. UK: Intermediate Technology Publications.

Stachowiak, S. (2009). *Pathways for change. 6 theories about how policy change happens*. Organizational Research Services (ORS).

Themistocleous, G., & Wearne, S. H. (2000). Project management topic coverage in journals. *International Journal of Project Management, 18* (1), 7–11.

Tough, A. M. (1967). *Learning without a teacher. A study of tasks and assistance during adult self-teaching projects*. Toronto: The Ontario Institute for Studies on Education (OISE).

Tough, A. M. (1968). *Why adults learn*. Toronto: The Ontario Institute for Studies on Education (OISE).

Tough, A. M. (1971). *The adult's learning projects. A fresh approach to theory and practice in adult learning* (2nd ed.). Austin, TX: Learning Concepts.

Travisan, M. S. (2002). Evaluation capacity in K-12 counseling programs. *American Journal of Evaluation, 23*(3), 275–290. http://dx.doi.org/10.1177/109821400202300304.

Ubels, J., Acquaye-Baddoo, N.-A., & Fowler, A. (Eds.) (2010). *Capacity development in practice*. Washington, DC: Earthscan.

UNDP (1997). *Capacity development resources book*. New York: UNDP.

UNDP (2009). *Overview of UNDP's approach to supporting capacity development*. Capacity Development Group, Bureau for Development Policy. United Nations Development Programme.

UNDP (2010). *Measuring capacity*. New York: United Nations Development Programme.

UNDP (2013). *Innovations in monitoring & evaluating results*. Discussion paper. Knowledge, innovation and capacity. United Nations Development Program.

UNESCO (2006). *Capacity building guidebook for planning education in emergencies and reconstruction*. Paris: International Institute for Educational Planning.

Valters, C. (2015). 3 big problems with how we think about results and development. *Global Views. The future of DFID*. Retrieved July 4, 2015, from <https://www.devex.com/news/3-big-problems-with-how-we-think-about-results-and-development-86419>.

van Kemenade, E. (2014). *The myth of the PDCA-cycle in times of emergent change*. Paper presented at the European Organization for Quality Congress, Goteborg, Germany.

Vogel, I. (2012). *Review of the use of 'theory of change' in international development*. London: DFID.

Watson, D. (2010). Measuring capacity development. Combining the "best of two worlds" in monitoring and evaluation of capacity development. In J. Ubels, N.-A. Acquaye-Baddoo, & A. Fowler (Eds.), *Capacity development in practice* (pp. 251–263). London: Earthscan.

Wehn, U. (2014a). *Effective knowledge and capacity development for enhancing the post-2015 development goals*. Paper presented at the International Development Studies Lecture Series 2014–2015. University of Amsterdam.

Wehn, U. (2014b). *Way forward. How can we strengthen institutional capacity?*. Delft, NL: UNESCO-IHE.

Wehn de Montalvo, U., & Alaerts, G. (2013). Leadership in knowledge and capacity development in the water sector: A status review. *Water Policy, 15*(Suppl. 2), 1–14. http://dx.doi.org/10.2166/wp.2013.109.

WRI (2008). *Roots of resilience. Growing the wealth of the poor*. Washington, DC: World Resources Institute (WRI), in collaboration with United Nations Development Programme, United Nations Environment Programme, and the World Bank.